



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 176049**

**TO: Ginny Portner  
Location: REM/3B02/3C18  
Art Unit: 1645  
Thursday, January 12, 2006  
Case Serial Number: 10/662126**

**From: Toby Port  
Location: Biotech-Chem Library  
REM-1A59  
Phone: 571-272-2523**

**[toby.port@uspto.gov](mailto:toby.port@uspto.gov)**

### **Search Notes**

**Examiner Portner,**

**See attached results.**

**If you have any questions about this search feel free to contact me at any time.**

**Thank you for using STIC search services!**

**Toby Port  
X22523**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**STIC-Biotech/ChemLib**

176049

---

**From:** Portner, Ginny  
**Sent:** Monday, January 09, 2006 4:32 PM  
**To:** STIC-Biotech/ChemLib  
**Subject:** 10/662,126

please search SEQ ID NO 36 and 37; oligomer search both sequences "at least 8 or 10 consecutive amino acids" and "at least 24 to 30 nucleotides". thanks

Ginny Portner  
Remsen Building  
Art Unit 1645  
Room E03, B02  
(571) 272-0862

RECEIVED  
JAN - 9 2006  
STIC

1/9/2006

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 00:34:19 ; Search time 3013 Seconds  
(without alignments)  
9055.711 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggaacaagccatcaaaa.....tgattgcagagagattaa 480

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sta:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vi:\*  
14: gb\_hlg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	480	100.0	480	6	BD009867	BD009867 Proteins,
2	480	100.0	13222	1	AE000637	AE000637 Helicobac
3	454.4	94.7	14404	1	AE001551	AE001551 Helicobac
4	449.6	93.7	480	1	AY725427	AY725427 Helicobac
5	294	61.3	853	6	AX788995	AX788995 Sequence
6	166	34.6	302212	1	AE017147	AE017147 Helicobac
7	165.2	34.4	347961	1	BX571657	BX571657 Wolinella
8	135.2	28.2	314150	1	CJ11168X1	AL139074 Campyloba
9	133.6	27.8	110000	1	CP000025_02	Continuation (3 of
10	113.8	23.7	300029	1	AE015936	AE015936 Clostridi
11	107.6	22.4	34083	1	AF124757	AF124757 Zymomonas
12	107.6	22.4	110000	1	AE008692_11	Continuation (12 o
13	106.4	22.2	110000	1	AE017180_24	Continuation (25 o
14	103.8	21.6	197615	5	BX470161	BX470161 Zebrafish
15	103.8	21.6	329100	1	SM5591787	AL591787 Snorhizo
16	101.4	21.1	6444	1	AF461795	AF461795 Bartonell
17	101.4	21.1	110000	1	BX897699_07	Continuation (8 of
18	100	20.8	659	6	AX789107	AX789107 Sequence

C 19	99.2	20.7	110000	1	BA000012_04	Continuation (5 of
C 20	99.2	20.7	110000	1	BA000012_05	Continuation (6 of
C 21	98	20.4	110000	1	BX897700_08	Continuation (9 of
C 22	97	20.2	110000	1	BA000040_53	Continuation (54 o
C 23	95.2	19.8	295250	1	BX321862	BX321862 Nitrosomo
C 24	95	19.8	110000	1	CP000083_16	Continuation (17 o
C 25	94.8	19.7	110000	1	CP000107_12	Continuation (13 o
C 26	94	19.6	300540	1	AE017154	AE017154 Haemophil
C 27	93.2	19.4	110000	1	AE017321_00	AE017321 Wolbachia
C 28	93	19.4	110000	1	CP000020_21	Continuation (22 o
C 29	92.8	19.3	302040	1	AE017317	AE017317 Desulfovi
C 30	92.6	19.3	10625	1	AE004296	AE004296 Vibrio ch
C 31	92	19.2	348526	1	CR378672	CR378672 Photobact
C 32	89.6	18.7	110000	1	AE016827_04	Continuation (5 of
C 33	88.8	18.5	110000	1	CR555306_35	Continuation (36 o
C 34	88.6	18.5	1755	6	AX414931	AX414931 Sequence
C 35	88.6	18.5	1927	6	AX416532	AX416532 Sequence
C 36	88.6	18.5	291091	1	AE017330	AE017330 Listeria
C 37	88.6	18.5	324050	1	AL591983	AL591983 Listeria
C 38	88.6	18.5	349980	6	AX641672	AX641672 Sequence
C 39	88.4	18.4	110000	1	AE014291_11	Continuation (12 o
C 40	88.2	18.4	11371	1	AE006236	AE006236 Pasteurel
C 41	88	18.3	347071	1	BX640415	BX640415 Bordetell
C 42	88	18.3	348997	1	BX640427	BX640427 Bordetell
C 43	88	18.3	349008	1	BX640444	BX640444 Bordetell
C 44	87	18.1	305050	1	AL596173	AL596173 Listeria
C 45	87	18.1	349980	6	AX417036	AX417036 Sequence

## ALIGNMENTS

RESULT 1	BD009867	480 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD009867	Proteins, in particular membrane proteins, of Helicobacter pylori,			
DEFINITION	BD009867	their preparation and use.			
ACCESSION	BD009867				
VERSION	BD009867.1	GI:18638240			
KEYWORDS	JP 2001502886-A/8.				
SOURCE	Helicobacter pylori				
ORGANISM	Helicobacter pylori				
REFERENCE	1	(bases 1 to 480)			
AUTHORS	Knaapp,B., Hundt,E. and Schmidt,K.H.				
TITLE	Proteins, in particular membrane proteins, of Helicobacter pylori,				
JOURNAL	their preparation and use				
COMMENT	Patent: JP 2001502886-A 8 06-MAR-2001;				
	CHIRON BEHRING GMBH & CO				
	OS Helicobacter pylori				
	PN JP 2001502886-A/8				
	PD 06-MAR-2001				
	PF 25-JUL-1997	JP 1998508651			
	PI 26-JUL-1996	DE 196 30 390.7			
	CJ2N15/31	C07K14/205, C07K16/12, G01N33/53, A61K31/70, A61K39/106, PC			
	CC Strandedness: Single;				
	CC Topology: Linear;				
	FT Key	Location/Qualifiers			
	FT source	1..480			
	FT	Location/Qualifiers			
	source	1..480			
	/organism="Helicobacter pylori"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:210"				
ORIGIN					
Query Match	100.0%;	Score 480;	DB 6;	Length 480;	
Best Local Similarity	100.0%;	Pred. No. 3.8e-116;			
Matches 480;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	



QY 1 ATGGAACAAAGCCATCAAACTTGAATCTCAATTTTATAGAGCATATCTTACAAATT 60  
Db 1 ATGGAACAAAGCCATCAAACTTGAATCTCAATTTTATAGAGCATATCTTACAAATT 60  
QY 61 CTACCTCACCGTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAGCCATAAA 120  
Db 61 CTACCTCACCGTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAGCCATAAA 120  
QY 121 AAAATTTGCTTTATAAGAAATATCACTTTTAATGAAGAGCTGTTTAACGGGCAATTCCT 180  
Db 121 AAAATTTGCTTTATAAGAAATATCACTTTTAATGAAGAGCTGTTTAACGGGCAATTCCT 180  
QY 181 AATAAGCCCAATTTCCCGGGCTTTTGATCTGTAGAGGCGATGCGCAAAACGGGAGGTTT 240  
Db 181 AATAAGCCCAATTTCCCGGGCTTTTGATCTGTAGAGGCGATGCGCAAAACGGGAGGTTT 240  
QY 241 TTAGCCTTCACTAGCTTTGGGGGTTTGAACCTCGAATCGCCAAACAAAAATCGTGTAT 300  
Db 241 TTAGCCTTCACTAGCTTTGGGGGTTTGAACCTCGAATCGCCAAACAAAAATCGTGTAT 300  
QY 301 TTTATGACCATGATGAAGCTTAAATTCCTGATCCCTGTAACCCGAGCGACAGATTAGAA 360  
Db 301 TTTATGACCATGATGAAGCTTAAATTCCTGATCCCTGTAACCCGAGCGACAGATTAGAA 360  
QY 361 TACCATTATAGAGTCTTAAAGCATAGGCGATGATCGGCAAGTGGTGCGACGGCTCAA 420  
Db 361 TACCATTATAGAGTCTTAAAGCATAGGCGATGATCGGCAAGTGGTGCGACGGCTCAA 420  
QY 421 GTGGATGGCAAGTGTCTGCTGAAGCCGATGAAAGCCATGATTCAGAGAGAGATTAA 480  
Db 421 GTGGATGGCAAGTGTCTGCTGAAGCCGATGAAAGCCATGATTCAGAGAGAGATTAA 480

RESULT 2  
AE000637/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

AE000637 13222 bp DNA linear BCT 06-APR-1999  
Helicobacter pylori 26695 section 115 of 134 of the complete genome.  
AE000637 AE000511  
AE000637.1 GI:2314536  
Helicobacter pylori 26695  
Helicobacter pylori 26695  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Helicobacteraceae; Helicobacter.  
1 (bases 1 to 13222)  
Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.  
The complete genome sequence of the gastric pathogen Helicobacter pylori  
Nature 388 (6642), 539-547 (1997)  
9252185  
2 (bases 1 to 13222)  
Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Wathey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.  
Direct Submission  
Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
3 (bases 1 to 13222)

White, O.  
Direct Submission  
Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
FEATURES  
source  
gene  
CDS  
gene  
CDS  
gene  
CDS  
gene  
CDS

1..13222  
/organism="Helicobacter pylori 26695"  
/mol\_type="genomic DNA"  
/strain="26695"  
/db\_xref="taxon:85962"  
complement(72..1343)  
/gene="HP1366"  
complement(72..1343)  
/gene="HP1366"  
/note="similar to SP:P23191 GB:X56977 PID:44182 percent identity: 37.14; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="type IIS restriction enzyme R protein (MBOIR)"  
/protein\_id="AAD08410.1"  
/db\_xref="GI:2314537"  
/translation="MPKLEKILLETITOLDPSKECKELANRIKSDYRGLHLSOHRNY DQNKITIQAFNEVGDFLOINTDMSKRPSNIGEEIYAKVVDNICKSEMPQDNS GKNQVQDSLRKKNLFVDMHMGILIERNNKPEPTNYIQSNIKYISLTPALAEFLNA QDLRKNFCYTOALENLQGFACREVMIEDNHYLDIEEMFVFTFNINFTSRSE IIEVVRGYSLSRIQEKELKELVODYCNPNHFNKLEKRDYHWNKNGAQOIFSLLEQ SVFEFTNKERLILKALNEENKNDKLRKSIKEKALYFEKHGKKEGKELHIIHVLPC LARSIEPDLDDKWNLIYIDAFNHAKISQTNKHICLYFKNCDDVLSKGLKEQBSL YLTIVNLYKLDNLAUKNKKDLSKNG"  
complement(1359..2141)  
/gene="HP1367"  
complement(1359..2141)  
/gene="HP1367"  
/note="similar to SP:P23192 GB:X56977 PID:44181 percent identity: 59.27; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="type IIS restriction enzyme M1 protein (mod)"  
/protein\_id="AAD08411.1"  
/db\_xref="GI:2314538"  
/translation="MILNKIYIEDVFTFLDKLEKSDVLAIDIPPYFNKIASWDSFKN DESPLTPSYAWIDKMLPKLQDTSFYFNTPFNCALFLAYLHKHVKHFLNFTVWKDD GFANAKRYNHAOSILFYSMHKQNTFNADIEIRIAYESAERIKAQSGIKLNKKRW FPNPKGLDLDWEITSQHVKEKGLKPKHPSIKPKALIERMIKASSHKNLDLID LFGSGMISLVAKSLERNFICGCEHAEYVHGSLEMFYNECE"  
complement(2128..2991)  
/gene="HP1368"  
complement(2128..2991)  
/gene="HP1368"  
/note="similar to SP:P29568 percent identity: 32.96; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="type IIS restriction enzyme M2 protein (mod)"  
/protein\_id="AAD08412.1"  
/db\_xref="GI:2314539"  
/translation="MNINKPYHSSSTNNNEVPDSDVLIITSPYFNKIDYANGTQD LQSAOHEVDIGALEYDYLGLKVMLECYRALKPNGLKLCINVLPMMLKVLNTH YNRHIFDLHADIQHSILHLDNNLNKPKMFLDLYIWKRNPTKRLMFGSYPRNF YAONTIEFIVFKDGPKEPTKEQKOSQTOBEWVEFTQIWEIPIPNKNDIAPGK HAALMPAELARLRLISCVGDDVLDPPSGSGTTLREAKLLKRNFTIGVLENYKPLI EQKLGUNLDFDE"  
3240..4664  
/gene="HP1369"  
/note="This region contains an authentic frame shift and is not the result of a sequencing artifact; similar to SP:P40814 percent identity: 39.74; identified by sequence similarity; putative; type III restriction enzyme M 4802..5668  
/gene="HP1370"

CDS

```
4802. .5668
/genes="HP1370"
/note="similar to GB:L24203 PID:1006312 PID:1221173
PID:1205306 PID:1574609 percent identity: 38.41;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="type III restriction enzyme M protein (mod)"
/protein_id="AAD08413.1"
/db_xref="GI:2314540"
/translation="MYKKLDENGFKYKLDPMVAKSGNNFSPYTFNLNGKTSWSPSGTFRW
YSGTLKDEQNNRIYFNGKNPMKRYLSEVAGRSKSTFWDSGVYNNLNGDAEIKR
LFGNKKVNNPKPEALLORILETSTKENDLVLDFAGSGTTCAVAHMKREKYIGIENG
EHPESVILRLKLVIGGFGSKAGKPDFGGVIVKYVELESYEEILRKLYEDNKKPLAY
DQYSDLVKRNESYTLINLEALEKMGVDIKETLENLWGVGFNFKEKVKVFRGNDEK
EILKALKEALIW"
5679. .8585
/genes="HP1371"
5679. .8585
/genes="HP1371"
/note="similar to GB:L25415 PID:496157 percent identity:
26.73; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="type III restriction enzyme R protein"
/protein_id="AAD08414.1"
/db_xref="GI:2314541"
/translation="MAKKKDLTDTDNEIFVAQKLAEEELNTEINEPLERLDFKSFNN
KELLDYQQOALINAPMLVAYFRDPKKNKKEFVAFVQKHYSFAHCDPAKKKLNPLKLS
HKFNHCYSPFENINELAFYATGSGKTIVILKVELLSVALRMGLIPKQIMFESA
NENLLOQFEKEIEKYNRNDYFPQIDFKLSKSVTHKDFYRAPKDSVIKQITLFIYRAD
LWNEDEENLKYKNDWNGENYVILDAHGNKSESRRQAFISLSLKGLFNFSA
TTTEESDLTAVNYGVGWKGLKESVLLKQNNLNAFKELKOLNREKREKIALKA
LILLGQKRYKTEYFVGLPMLVFTHSVNVKNSDAEIPFKTLARVLENDGDFLAK
EDLLELKNPELPSDDKDKVVKVFEKGLSKMPKGLKEVFYANNCHIIVINPK
NNQETAPKLTSDKVPCLIRIGDITWICEKLSKVYVSKNLSFKRESEFSQIDKSSI
NLVGSRTDQTSRPSVILFNLIGLDDAKKLVQKSGVRIEVSNNORQLAY
LDIDGAIKALPNAAMLETPIPTNYASLAILFKQESKNGRSGSWEIKLEK
TPIKHALFVPCRKQTSILELPEFASFKWSEKFNKDLKELNYFNLMSEKHFILKHEIYD
PKYVMOLKQTOBAPNKVTSWHYKOLDYMSIILKGLVPNOKVPKDBFNALDSKIV
HKPRLKVKADKKEELVLTQEVKEVAPLDKETLILKLAQGEIDPDYDTKHKONKTFK
GSAELLKLEHYTPIILKAKNDWLKHVVVKESESDPLELLIKIETLEQNTDFWAFS
KIDEHLNLFIFYNNAERKFPDFIFWLEKGGTQIICFIDPKSGSKHYDIBHKADAY
QIFDKIKFNPDKNPLKIKVILKFKYGDKEVADGYRDIYWKKGLEDFFLKQLA"
/complement(8635..9381)
/genes="HP1372"
complement(8635..9381)
/genes="HP1372"
/note="similar to GB:L42023 SP:P44475 PID:1004176
PID:1221943 PID:1204296 percent identity: 33.56;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreC)"
/protein_id="AAD08415.1"
/db_xref="GI:2314542"
/translation="WRFFYFLMLLIGLIFVFLDPKGSSTVSDIRKKNALNKNLSL
LDNVQAYFQQAQNIKEFKERLILKALKNLADKLERLNSIYPLENPKMTYTFPMTS
FISLSDTSVNLNINLEENKILGLVNSAIGIAVLEKGRINGFLNAHKRCAYSM
IGQNVGLFIGNFKOELVVDVFPVSAEINIGDQVLTSLGIDFGAGVFGVSVSIED
HYTKSAVLKNAFLSKALRHVFLSDVKN"
/complement(9385..10428)
/genes="HP1373"
complement(9385..10428)
/genes="HP1373"
/note="similar to SP:P13519 GB:M22055 PID:606191
PID:1128967 GB:U00096 percent identity: 51.87; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreB)"
/protein_id="AAD08416.1"
/db_xref="GI:2314543"
```

gene

CDS

```
5679. .8585
/genes="HP1371"
5679. .8585
/genes="HP1371"
/note="similar to GB:L25415 PID:496157 percent identity:
26.73; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="type III restriction enzyme R protein"
/protein_id="AAD08414.1"
/db_xref="GI:2314541"
/translation="MAKKKDLTDTDNEIFVAQKLAEEELNTEINEPLERLDFKSFNN
KELLDYQQOALINAPMLVAYFRDPKKNKKEFVAFVQKHYSFAHCDPAKKKLNPLKLS
HKFNHCYSPFENINELAFYATGSGKTIVILKVELLSVALRMGLIPKQIMFESA
NENLLOQFEKEIEKYNRNDYFPQIDFKLSKSVTHKDFYRAPKDSVIKQITLFIYRAD
LWNEDEENLKYKNDWNGENYVILDAHGNKSESRRQAFISLSLKGLFNFSA
TTTEESDLTAVNYGVGWKGLKESVLLKQNNLNAFKELKOLNREKREKIALKA
LILLGQKRYKTEYFVGLPMLVFTHSVNVKNSDAEIPFKTLARVLENDGDFLAK
EDLLELKNPELPSDDKDKVVKVFEKGLSKMPKGLKEVFYANNCHIIVINPK
NNQETAPKLTSDKVPCLIRIGDITWICEKLSKVYVSKNLSFKRESEFSQIDKSSI
NLVGSRTDQTSRPSVILFNLIGLDDAKKLVQKSGVRIEVSNNORQLAY
LDIDGAIKALPNAAMLETPIPTNYASLAILFKQESKNGRSGSWEIKLEK
TPIKHALFVPCRKQTSILELPEFASFKWSEKFNKDLKELNYFNLMSEKHFILKHEIYD
PKYVMOLKQTOBAPNKVTSWHYKOLDYMSIILKGLVPNOKVPKDBFNALDSKIV
HKPRLKVKADKKEELVLTQEVKEVAPLDKETLILKLAQGEIDPDYDTKHKONKTFK
GSAELLKLEHYTPIILKAKNDWLKHVVVKESESDPLELLIKIETLEQNTDFWAFS
KIDEHLNLFIFYNNAERKFPDFIFWLEKGGTQIICFIDPKSGSKHYDIBHKADAY
QIFDKIKFNPDKNPLKIKVILKFKYGDKEVADGYRDIYWKKGLEDFFLKQLA"
/complement(8635..9381)
/genes="HP1372"
complement(8635..9381)
/genes="HP1372"
/note="similar to GB:L42023 SP:P44475 PID:1004176
PID:1221943 PID:1204296 percent identity: 33.56;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreC)"
/protein_id="AAD08415.1"
/db_xref="GI:2314542"
/translation="WRFFYFLMLLIGLIFVFLDPKGSSTVSDIRKKNALNKNLSL
LDNVQAYFQQAQNIKEFKERLILKALKNLADKLERLNSIYPLENPKMTYTFPMTS
FISLSDTSVNLNINLEENKILGLVNSAIGIAVLEKGRINGFLNAHKRCAYSM
IGQNVGLFIGNFKOELVVDVFPVSAEINIGDQVLTSLGIDFGAGVFGVSVSIED
HYTKSAVLKNAFLSKALRHVFLSDVKN"
/complement(9385..10428)
/genes="HP1373"
complement(9385..10428)
/genes="HP1373"
/note="similar to SP:P13519 GB:M22055 PID:606191
PID:1128967 GB:U00096 percent identity: 51.87; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreB)"
/protein_id="AAD08416.1"
/db_xref="GI:2314543"
```

gene

CDS

```
complement(8635..9381)
/genes="HP1372"
complement(8635..9381)
/genes="HP1372"
/note="similar to GB:L42023 SP:P44475 PID:1004176
PID:1221943 PID:1204296 percent identity: 33.56;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreC)"
/protein_id="AAD08415.1"
/db_xref="GI:2314542"
/translation="WRFFYFLMLLIGLIFVFLDPKGSSTVSDIRKKNALNKNLSL
LDNVQAYFQQAQNIKEFKERLILKALKNLADKLERLNSIYPLENPKMTYTFPMTS
FISLSDTSVNLNINLEENKILGLVNSAIGIAVLEKGRINGFLNAHKRCAYSM
IGQNVGLFIGNFKOELVVDVFPVSAEINIGDQVLTSLGIDFGAGVFGVSVSIED
HYTKSAVLKNAFLSKALRHVFLSDVKN"
/complement(9385..10428)
/genes="HP1373"
complement(9385..10428)
/genes="HP1373"
/note="similar to SP:P13519 GB:M22055 PID:606191
PID:1128967 GB:U00096 percent identity: 51.87; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreB)"
/protein_id="AAD08416.1"
/db_xref="GI:2314543"
```

gene

CDS

```
complement(8635..9381)
/genes="HP1372"
complement(8635..9381)
/genes="HP1372"
/note="similar to GB:L42023 SP:P44475 PID:1004176
PID:1221943 PID:1204296 percent identity: 33.56;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreC)"
/protein_id="AAD08415.1"
/db_xref="GI:2314542"
/translation="WRFFYFLMLLIGLIFVFLDPKGSSTVSDIRKKNALNKNLSL
LDNVQAYFQQAQNIKEFKERLILKALKNLADKLERLNSIYPLENPKMTYTFPMTS
FISLSDTSVNLNINLEENKILGLVNSAIGIAVLEKGRINGFLNAHKRCAYSM
IGQNVGLFIGNFKOELVVDVFPVSAEINIGDQVLTSLGIDFGAGVFGVSVSIED
HYTKSAVLKNAFLSKALRHVFLSDVKN"
/complement(9385..10428)
/genes="HP1373"
complement(9385..10428)
/genes="HP1373"
/note="similar to SP:P13519 GB:M22055 PID:606191
PID:1128967 GB:U00096 percent identity: 51.87; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreB)"
/protein_id="AAD08416.1"
/db_xref="GI:2314543"
```

gene

CDS

```
complement(8635..9381)
/genes="HP1372"
complement(8635..9381)
/genes="HP1372"
/note="similar to GB:L42023 SP:P44475 PID:1004176
PID:1221943 PID:1204296 percent identity: 33.56;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreC)"
/protein_id="AAD08415.1"
/db_xref="GI:2314542"
/translation="WRFFYFLMLLIGLIFVFLDPKGSSTVSDIRKKNALNKNLSL
LDNVQAYFQQAQNIKEFKERLILKALKNLADKLERLNSIYPLENPKMTYTFPMTS
FISLSDTSVNLNINLEENKILGLVNSAIGIAVLEKGRINGFLNAHKRCAYSM
IGQNVGLFIGNFKOELVVDVFPVSAEINIGDQVLTSLGIDFGAGVFGVSVSIED
HYTKSAVLKNAFLSKALRHVFLSDVKN"
/complement(9385..10428)
/genes="HP1373"
complement(9385..10428)
/genes="HP1373"
/note="similar to SP:P13519 GB:M22055 PID:606191
PID:1128967 GB:U00096 percent identity: 51.87; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreB)"
/protein_id="AAD08416.1"
/db_xref="GI:2314543"
```

```
/translation="WIFSKLIGLFSHDIAIDLTANTIVLVKGGQIIINPSIVAVRM
GLPDSKAYDILAVGSEAKEMLGKTPNSRAIRPKDGVIAVDITAKMIRPIEKVHK
RKTWIRPRIMVCFYGLTSVERNAVKESTLSAGAREVFLIEPMAAIGAGLPVKEPQ
GSLIVDIGGGTTEIGVLSGLVSKISRVAGDKLDQSDIVIRYKFKNLLIGERTGE
IKTEIGCAIKLDPPLTMEVSGRDQVSGLLHTIELSSDDVFAIKDQVREISALRSVL
BEVKPDLAKDIOVQGVVLITGGGALIKGLDKVLSDMVLPVYVVDPEPLLAAGKTGEAI
QDLDLLSRVGFSE"
complement(10471..11811)
/genes="HP1374"
complement(10471..11811)
/genes="HP1374"
/note="similar to GB:L18867 SP:P3138 GB:Z23278 PID:347729
PID:388977 percent identity: 55.04; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP-dependent protease ATPase subunit (c1px)"
Query Match 100.0%; Score 480; DB 1; Length 13222;
Best Local Similarity 100.0%; Pred. No. 3.3e-116;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAAGCCCATCAAACTTCGAATCTCAATTTTATAGAGCATATCTTACAAATT 60
DB 13108 ATGGAACAAAGCCCATCAAACTTCGAATCTCAATTTTATAGAGCATATCTTACAAATT 13049
QY 61 CTACCTCACCGCTATCCATGCTTTTATAGTATAGAAATTTATAGAGTTTACAAGCCCAATAA 120
DB 13048 CTACCTCACCGCTATCCATGCTTTTATAGTATAGAAATTTATAGAGTTTACAAGCCCAATAA 12989
QY 121 AAAATTGCTGCTTATAGAAATATATCACTTTTATGAAGACGTTGTTAAACGGGCAATTTCCT 180
DB 12988 AAAATTGCTGCTTATAGAAATATATCACTTTTATGAAGACGTTGTTAAACGGGCAATTTCCT 12929
QY 181 AATAAGCCCAATTTTCCCGGGGTTTTCATCTGATAGAGGCGATGCGCAACCGGAGGTTT 240
DB 12928 AATAAGCCCAATTTTCCCGGGGTTTTCATCTGATAGAGGCGATGCGCAACCGGAGGTTT 12869
QY 241 TTAGCCCTTCACTAGCTGTGGGGTGTGACCTCAAAATCGCAAAACAAAAATCTGTAT 300
DB 12868 TTAGCCCTTCACTAGCTGTGGGGTGTGACCTCAAAATCGCAAAACAAAAATCTGTAT 12809
QY 301 TTATGACGAGTTGATAGGTTAAATTCGATCCCTGTAAACCCGAGCGACAGATTAGAA 360
DB 12808 TTATGACGAGTTGATAGGTTAAATTCGATCCCTGTAAACCCGAGCGACAGATTAGAA 12749
QY 361 TACCATTTAGAACTTTAAAGCATAAAGGCGATGATGCGCAAGTGGTGGCACGGCTCAA 420
DB 12748 TACCATTTAGAACTTTAAAGCATAAAGGCGATGATGCGCAAGTGGTGGCACGGCTCAA 12689
QY 421 GTGGATGGCAAGTGGTGGTGAAGCCGAATTGAAAGCCATGATTGCAAGAGAGATTAA 480
DB 12688 GTGGATGGCAAGTGGTGGTGAAGCCGAATTGAAAGCCATGATTGCAAGAGAGATTAA 12629
```

```
RESULT 3
AE001551/c
LOCUS Helicobacter pylori, strain J99 section 112 of 132 of the complete genome.
DEFINITION
ACCESSION AE001551
VERSION AE001551.1
KEYWORDS GI:4155887
SOURCE Helicobacter pylori J99
ORGANISM Helicobacter pylori J99
REFERENCE 1 (bases 1 to 14404)
AUTHORS Alm,R.A., Ling,L.S., Moir,D.T., King,B.L., Brown,E.D., Doig,P.C., Smith,D.R., Noonan,B., Guild,B.C., deJonghe,B.B., Carmel,G., Tummino,P.J., Caruso,A., Uria-Nickelsen,M., Mills,D.M., Ives,C., Gibson,R., Merberg,D., Mills,S.D., Jiang,Q., Taylor,D.E., Vovis,G.F. and Trust,T.J.
```

```
AE001551
LOCUS Helicobacter pylori, strain J99 section 112 of 132 of the complete genome.
DEFINITION
ACCESSION AE001551
VERSION AE001551.1
KEYWORDS GI:4155887
SOURCE Helicobacter pylori J99
ORGANISM Helicobacter pylori J99
REFERENCE 1 (bases 1 to 14404)
AUTHORS Alm,R.A., Ling,L.S., Moir,D.T., King,B.L., Brown,E.D., Doig,P.C., Smith,D.R., Noonan,B., Guild,B.C., deJonghe,B.B., Carmel,G., Tummino,P.J., Caruso,A., Uria-Nickelsen,M., Mills,D.M., Ives,C., Gibson,R., Merberg,D., Mills,S.D., Jiang,Q., Taylor,D.E., Vovis,G.F. and Trust,T.J.
```

```
AE001551
LOCUS Helicobacter pylori, strain J99 section 112 of 132 of the complete genome.
DEFINITION
ACCESSION AE001551
VERSION AE001551.1
KEYWORDS GI:4155887
SOURCE Helicobacter pylori J99
ORGANISM Helicobacter pylori J99
REFERENCE 1 (bases 1 to 14404)
AUTHORS Alm,R.A., Ling,L.S., Moir,D.T., King,B.L., Brown,E.D., Doig,P.C., Smith,D.R., Noonan,B., Guild,B.C., deJonghe,B.B., Carmel,G., Tummino,P.J., Caruso,A., Uria-Nickelsen,M., Mills,D.M., Ives,C., Gibson,R., Merberg,D., Mills,S.D., Jiang,Q., Taylor,D.E., Vovis,G.F. and Trust,T.J.
```

```
AE001551
LOCUS Helicobacter pylori, strain J99 section 112 of 132 of the complete genome.
DEFINITION
ACCESSION AE001551
VERSION AE001551.1
KEYWORDS GI:4155887
SOURCE Helicobacter pylori J99
ORGANISM Helicobacter pylori J99
REFERENCE 1 (bases 1 to 14404)
AUTHORS Alm,R.A., Ling,L.S., Moir,D.T., King,B.L., Brown,E.D., Doig,P.C., Smith,D.R., Noonan,B., Guild,B.C., deJonghe,B.B., Carmel,G., Tummino,P.J., Caruso,A., Uria-Nickelsen,M., Mills,D.M., Ives,C., Gibson,R., Merberg,D., Mills,S.D., Jiang,Q., Taylor,D.E., Vovis,G.F. and Trust,T.J.
```

TITLE Genomic-sequence comparison of two unrelated isolates of the human  
gastric pathogen *Helicobacter pylori*  
JOURNAL Nature 397 (6715), 176-180 (1999)  
PUBMED 9923682  
REFERENCE 2 (bases 1 to 14404)  
AUTHORS King,B.L., Alm,R.A. and Trust,T.J.  
TITLE Direct Submission  
JOURNAL Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney  
Street, Cambridge, MA 02139, USA  
COMMENT Address all correspondence to: hpc@arcb.us.astra.com or Richard  
A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge,  
MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith,  
Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills,  
Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics  
Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and  
Diane E. Taylor are with the University of Alberta Department of  
Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G  
2H7 and the Canadian Bacterial Diseases Network. All other authors  
are with Astra Research Center Boston, 128 Sidney Street,  
Cambridge, MA, 02139. Putative identifications, sequence  
alignments, and name and sequence search capability are available  
at ARCB's World Wide Web site. (URL:  
http://www.astra-boston.com/hpylori).

FEATURES  
source

1. .14404  
/organism="Helicobacter pylori J99"  
/mol\_type="genomic DNA"  
/strain="J99"  
/db\_xref="taxon:85963"  
complement (71. .1531)  
/gene="dnaB"  
/note="synonym: jhp1280"  
complement (71. .1531)  
/gene="dnaB"  
/note="similar to H. pylori 26695 gene HP1362"  
/codon\_start=1  
/transl\_table=11  
/product="REPLICATIVE DNA HELICASE"  
/protein\_id="AAD06865.1"  
/db\_xref="GI:4155895"

/translation="MDHLKHLQQLQNIERIVLSGIVLANHKIEHSVLEPSDFYPP  
HGLFEALKHEVNCIPIDENIRQMPKQKQISEDDLVAIFAAPIDNIEAYEISIK  
NASIKRLTLANTREQALESAQSDILNAVEREYVALLNGSTTEGFRGKEVLES  
TNWLTENQKSLTKVTGPTGCVQDNTVTSFGNQSIVLGLARPMSGKTSLMNMVL  
SALNDRGVAVFSLEASQALRALSDLTSLNMHDELSARLDDOENLAKCFDHL  
QKLLPFYDKSVYRMDQIRLQRLKLSQHKELGAFIDYQLQMSGNATKERHQIAEI  
SHEKTLARELEIPIIALVOLNRSLRDRDKPILSDIKDSGGIQDQADIVLFLRGY  
IYQRAEDNKIDIKKEGVFEAAQELHLKNEERIRHKONGSTEEAEIIVAKNRNGAT  
GTVYTRFNAPTFRYEDMPVDLSHEGQETKFEMPTT"

complement (1542. .2936)  
/gene="jhp1281"  
complement (1542. .2936)  
/gene="jhp1281"  
/note="similar to H. pylori 26695 gene HP1363"  
/codon\_start=1  
/transl\_table=11  
/product="putative"

/protein\_id="AAD06866.1"  
/db\_xref="GI:4155896"  
/translation="MLSVTEKGNALDKRVLEWLLSIEDILMENAAMALERAVIQNASL  
GARVILICSGONGGPGYTLARLVGRFKTLVPFMKLAKSPMCQLOKERAKVGVVIK  
ANEEKNEDECVDLVDCVVGSAFKGLEPDPESISOKARFKIACDIPSGIDSKGRV  
DKRAFADTTISGAILKSLLSKADYIGELKVHGVNQIYVPTDPTLEKSDL  
KIPLRKNAHKDYGHAHVLLGSGAGLLSALSFGSVVYSIQALECEITISNNK  
VLELTCFEPFKLSAAGNGLNTPKDFKWELEAPCVDLGVFHKVLEQALKEKE  
PILVTHPEKFLSLKSGVGINISMLELDNKLKLEIARDPSQKYPKVVLKLGANTLIHQ  
GRVFINLGSVALAKAGSGDVLGLVLSLQNTYPLDAAINASLAHALAGLEFKNHY  
ALTPDLLEIKEL"

complement (2939. .3973)  
/gene="jhp1282"  
complement (2939. .3973)  
/gene="jhp1282"  
/note="similar to H. pylori 26695 gene HP1364"

/codon\_start=1  
/transl\_table=11  
/product="putative histidine kinase sensor protein"  
/protein\_id="AAD06867.1"  
/db\_xref="GI:4155897"  
/translation="MASKIASEIQLHMOHADYHNAALNALSISRYKDASTALPDSKKR  
VLYNVPESADLIKTHKEGFFSGEYILFSDETFAHIGVAKMLFKNSKFLHPSSLY  
RNMVLVFAVFLCVIGSVFLGRFLKIRNEIIRIDHFDLTKTTHETLPMALVSL  
KTELDQOHRRIKAIQRMSFLYSLSYVMQDIERESFALLDLKALITKENTLIFSEM  
IDYHLEFSDLVGVVLKAKQODFISLYSNLMAAIKYSVMHGVIHIELTPEFLKVK  
LGVEIPDKITTKLSVRYARNTSSVIGVIGLDLVKKYCEKVKMRLEIHSBSPKSGPH  
ENSFCVQFQG"  
complement (4107. .4748)  
/note="similar to H. pylori 26695 gene HP1365"  
/codon\_start=1  
/transl\_table=11  
/product="putative TRANSCRIPTIONAL REGULATOR"  
/protein\_id="AAD06868.1"  
/db\_xref="GI:4165002"  
/translation="MQKIFLEDDYLLSSIEKIEFHLGYEVCFAFNGKEAHERLSV  
ERFNLLDVQVPMNSLELFKIRKNDFLISTPVIFITALQDNATLKNAFLNGASDYL  
KKPFDLDELEARIKRFNDPDIETMNPFIYQHALNVKKEIILAPTAQLLFVLEH  
KGQIISQALENNLWEQAIDDSLTRTVIKVLRKLKGNKCIETHKGVGVRFNPL"  
complement (4110. .4748)  
/gene="jhp1283"  
4857. .6806  
/gene="jhp1284"  
4857. .6806  
/gene="jhp1284"  
/codon\_start=1  
/transl\_table=11  
/product="TYPE II DNA MODIFICATION ENZYME  
(METHYLTRANSFERASE)"  
/protein\_id="AAD06858.1"  
/db\_xref="GI:4155888"  
/translation="MKSETIYKQPCFELGILFNFTSLLESKKYNEKVDLIFDLKSI  
VKDQTKNTLNFSTLSSKGTQTKNEILKECSNQVKLDEETLKXAFVFKQGSW  
DYPIHKNQOFLKEQDLVLFYLFKEMTAFDHKRLNGINTIKVLEVALVSEFEN  
ELCIHKNPRLVNLNHSPTVSLDKAKNYDLSKITNHPNYPKQVKEWODLNLADNL  
LGNEFLPDLTVFKDLSEEVKSLFNEENGLTIKSENYQALNSLKRYKEAIDCIYI  
DPFNTGSDPAYIDRFQDSTWSLMNRLQLAVDFLSPOGNFYHLHRYANLYGEMLL  
NDIFSKENFNEIILWHFRYTGQIQSNFPRKSHLSLWYSCNWNVNFRTYSDNYK  
TVYDRWREFIVDNKIVYPNYPKADSRFDYLRLOSTKEPKNGDIIATINGVID  
DWTVDIAIDPKADLERLOGLTQKPEKLERIISKASNSNISVCDFFAGSGTTCAVA  
HKLKRYIVGMEGEHPESVILPRLKVIGGFKSGALKBFNGGGYIKVVELESYEBILR  
KIKYENDKPLAYEEOYSLVERKNESYTLNIEALENGVDIKETLENLHGVGVFEFN  
EKVVKFGNDKESVILKALKALILW"  
6816. .9734  
/gene="jhp1285"  
6816. .9734  
/gene="jhp1285"  
/note="similar to H. pylori 26695 gene HP1371"  
/codon\_start=1  
/transl\_table=11  
/product="putative"  
/protein\_id="AAD06859.1"  
/db\_xref="GI:4155889"  
/translation="MAKKQEVNNEIFVQAQKLAEEELNANEINDPLEMDPKSPDDN  
KELLDYQQQALINAFRVLYVAFDFKESKKEFYAQKHSFANCDFAKKLNLHLKS  
HFYKHCVEKEIENRNFYMATGSGKTIIVIKVELLSVAMGGLIPKKNIMPFSA  
NEHLIKQFEKEIKYRNKDYSKQIDFNKLSVKNKDIFYAPKOSLMKEIALFYRAD  
LMSDEESKNLNYKDCWDNGENVVILDEAHKGNKTESKQAIIFSLISLKGFLNPSA  
TYTEESDLITAVNLSYGVGKESVLLKNNLNAPELKDNDKEKIALLKA  
LLLLGQMKRYKVEGYFHDPLMLVTHSVNNSDAEIEFETLARVIENDSDSPSKAK  
DDBLEUNKEFLFSDGKQKQKQYKIVFKEKSGMDFKGLKEAVFYASNGHIEVIN  
PKNNKEITAFNLSDFKICLIRIDGITETWIREKLKSVKWSKLSFKEESYFDKDS  
SINILVSGRAFDFDGTSPSVILFNLIGLDDAKLVKQSFGRGVRIESYKQNRQL  
AYLEIDRAIKDKLNAAMLEMLFVPTNHASLEAILKFOKESNGENGRGPPWRIKL  
EKTRIEHALFVPCYKQETWALKISQASFKWSEKFKDLKEHFLHMEKHFILKHEI  
YDPKAYLLKEMIQTAFKKVSTWHYKOLDYMISEIKGLYPNQKVPKDFNADNEK  
IVHFRYKVKADKKEKLIQTIQEVKEHAPLDKETLRKKIAQGEIDPYDETKHQDRTF  
KVGDAELLKKEHYTTLIKAKDCDWLKVVKVKSSEDFLEELKILITETLQENYDFA  
FSKIDEHLNLFIPYIDNATERRFPDFIFWLQKGGTQIICFIIDPKGSKHTDYEHKAD

gene

gene

CDS

gene

CDS

gene	AYOLPEDKVFNPCKDDPNLKIKVVLKPVGDKDDVGERYRDLWIEKGLKLEYFFSGFKGLA	
	complement (9784. .10530)	
CDS	/gene="jhp1286"	
	complement (9784. .10530)	
	/gene="jhp1286"	
	note="similar to H. pylori 26695 gene HP1372"	
	/codon_start=1	
	/transl_table=11	
	/product="putative"	
	/protein_id="AAD06860.1"	
	/db_xref="GI:4155890"	
	translation="MRFEYKFWLLGILFIYFLDIKSSSYISDRVKSALMSAKNSL	
	LDNVQAYFQONIKRFXERKILERLKLADLKERLNSIYPLENPKTYTPFTFMTS	
	FINLEDTHSVLNPIVNLNENKIYGLVSHQDAIGIAVLKRLGFLNNAHKRCAYSM	
	IQGNQVLGFTGNFKQELVDFVPSAEINIQDVLTSGLDGI FGAGVFVGVSSVED	
	HYTKSAVLKNAPLSEAKLRHVFSLGVKN"	
	complement (10534. .11577)	
	/gene="nureB"	
	/note="synonym: jhp1287"	
	complement (10534. .11577)	
	/gene="mreB"	
	note="similar to H. pylori 26695 gene HP1373"	
	/codon_start=1	
	/transl_table=11	
	/product="ROD SHAPE-DETERMINING PROTEIN"	
	/protein_id="AAD06861.1"	
	/db_xref="GI:4155891"	
	translation="MFSLKIGLGHSDIADLTANTIVLVKGGIILNEPSIVAVRM	
	GLFDSKAVDILAVGSEAKMLKTPNSIRAIRPKDGVIAVDITAKMIRYPIEKAKH	
	RKTWRPRIMVCVPIGLTSVERNVAKESAGAREVFLIEPMAAIGAGVVKPEPQ	
	GSLIVDIGGTEIGVI SLGGIVISKIRVAGDKLDQSIYIRKPNLLIGERTGEE	
	IKETIGCAIKLDPPLTMEVGRDQVSLHTIELSSDDVFEAIKDQVREISSALRSVL	
	BEVKDPLVRDVIQONGVLTGGGALIKGLDKYLSLVKPLPVYVGDPELLAVAKGTGEAI	
	QDLDLLSRVGFE"	
	complement (11630. .12988)	
	/gene="clpX"	
	/note="synonym: jhp1288"	
	94.7%; Score 454.4; DB 1; Length 14404;	
	Query Match	
	Best Local Similarity 96.7%; Pred. No. 2e-109; Indels 0; Gaps 0;	
	Matches 464; Conservative 0; Mismatches 16; Indels 0; Gaps 0;	
Qy	1 ATGGAACAAAGCCATCAAAACTTGCATCTCAATCTCAATTTTTATAGAGCATATCTTACAAATT 60	
Db	14284 ATGGAACAAAGCCATCAAAACTTCAATCTCAATTTTTATAGAGCATATCTTACAGATT 14225	
Qy	61 CTACCTCACCGGTATCCCATGCTTTTAGTGGATAGAAATTATAGATTACAAGCCCAATAAA 120	
Db	14224 TTACCTCACCGGTATCCCATGCTTTTAGTGGATAGAAATTATAGATTACAAGCCCAATAAA 14165	
Qy	121 AAAATTGTCGTTAAGAATATCACTTTTATAGAGCATGTTTAAACGGGCATTTCCCT 180	
Db	14164 AAAATTGTCGTTAAGAATATCACTTTTATAGAGCATGTTTAAACGGGCATTTCCCT 14105	
Qy	181 AATAAGCCCATTTCCCGGGCTTTTGATCGTAGAGGCATGGCCCAACGGGAGGTTT 240	
Db	14104 AATAAGCCCATTTCCCGGGCTTTTGATCGTAGAGGCATGGCCCAACGGGAGGTTT 14045	
Qy	241 TTACGCTTCACTAGCTTGGGGGGTTTGACCTCGAAATCGCCAAACAAAATCGTGAT 300	
Db	14044 TTACGCTTCACTAGCTTGGGGGGTTTGACCTCGAAATCGCCAAACAAAATCGTGAT 13985	
Qy	301 TTCATGAGCATTTAGAGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 360	
Db	13984 TTCATGAGCATTTAGAGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 13925	
Qy	361 TACCATTTAGAGTCTTAAAGCATAAAGGCGATGATCTGCGAAGTGGGGCGCGCTCAA 420	
Db	13824 TACCATTTAGAGTCTTAAAGCATAAAGGCGATGATCTGCGAAGTGGGGCGCGCTCAA 13865	
Qy	421 GTGGATGSCAAAGTGGTGGCTGAAGCCGAAATTGAAGAGCCATGATTGACAGAGAGATTAA 480	

Db	13864 GTGGATGSCAAAGTGGTGGCTGAAGCCGAAATTAAAGCCATGATCGCAGAGAGATAA 13805	
	RESULT 4	
LOCUS	AY725427	
	480 bp DNA linear BCT 08-JUL-2005	
DEFINITION	Helicobacter pylori (3R)-hydroxymyristoyl-acyl carrier protein	
	dehydratase (fabZ) gene, complete cds.	
ACCESSION	AY725427	
	GI:56684724	
KEYWORDS	Helicobacter pylori	
	Helicobacter pylori	
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;	
	Helicobacteraceae; Helicobacter.	
REFERENCE	1 (bases 1 to 480)	
	Liu, W., Luo, C., Han, C., Peng, S., Yang, Y., Yue, J., Shen, X. and	
AUTHORS	Jiang, H.	
	A new beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from	
TITLE	Helicobacter pylori: Molecular cloning, enzymatic characterization,	
	and structural modeling	
JOURNAL	Biochem. Biophys. Res. Commun. 333 (4), 1078-1086 (2005)	
	15967411	
PUBMED	2 (bases 1 to 480)	
	Liu, W., Han, C., Wang, Q., Jiang, H. and Shen, X.	
AUTHORS	Direct Submission	
	Submitted (16-AUG-2004) Drug Discovery and Design Center, Shanghai	
JOURNAL	Institute of Materia Medica, Shanghai Institutes for Biological	
	Sciences, Chinese Academy of Sciences, Zuchongzhi Road 555,	
	Shanghai 201203, China	
	Location/Qualifiers	
FEATURES	source	
	1. .480	
	/organism="Helicobacter pylori"	
	/mol_type="genomic DNA"	
	/strain="SS1"	
	/db_xref="taxon:210"	
gene	1. .480	
	/genes="fabZ"	
CDS	1. .480	
	/gene="fabZ"	
	/function="fatty acid and phosphatidic acid biosynthesis"	
	/note="enzyme"	
	/codon_start=1	
	/transl_table=11	
	/product="(3R)-hydroxymyristoyl-acyl carrier protein	
	dehydratase"	
	/protein_id="AAW22049.1"	
	/db_xref="GI:56684725"	
	translation="MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRITELQANQKIVA	
	YKNIIFNEVDVFNHGFNPKPIFGVLIVEGMAQSGFLAFTSLWGFDFEIAKTIKIVFM	
	TIDKVKFERIPVTPGDRLEHVLEVLKHGMIWQVGGTAQVDGKVAEELKAMIAERE"	
	ORIGIN	
Qy	Query Match	
	Best Local Similarity 93.7%; Score 449.6; DB 1; Length 480;	
Db	Matches 461; Conservative 0; Mismatches 19; Indels 0; Gaps 0;	
	1 ATGGAACAAAGCCATCAAAACTTGCATCTCAATTTTTATAGAGCATATCTTACAAATT 60	
Qy		
	1 ATGGAACAAAGCCATCAAAACTTGCATCTCAATTTTTATAGAGCATATCTTACAGATT 60	
Db		
	61 CTACCTCACCGGTATCCCATGCTTTTAGTGATAGATTATAGATTACAGCCCAATAAA 120	
Qy		
	61 TTACCTCACCGGTATCCCATGCTTTTAGTGATAGATTATAGATTACAGCCCAATAAA 120	
Db		
	121 AAAATTGTCGTTAAGAATATCACTTTTAAATCAAGACCGTGTTTTAAACGGGCGATTTCCCT 180	
Qy		
	121 AAAATTGTCGTTAAGAATATCACTTTTAAATCAAGACCGTGTTTTAAATGGCATTTCCCT 180	
Db		
	181 AATAAGCCCATTTCCCGGGCTTTTGATCGTAGAGGCATGGCCAAACGGGAGGTTT 240	
Qy		
	181 AATAAGCCCATTTCCCGGGCTTTTGATCGTAGAGGCATGGCCAAACGGGAGGTTT 240	
Db		

```
QY 241 TTAGCTTCACCTAGCTTGTGGGGTTTGACCTGAAATCGCAATCGCAAAACAAAAATCGTGAT 300
Db 241 TTGGCTTTCACCTAGCTTGTGGGGTTTGACCTGAAATAGCAAAACAAAAATCGTGAT 300
QY 301 TTCATGACGATTGATAAGCTTAAATCCGATCCCTGTAACCCAGCCAGCCAGACAGATTAGAA 360
Db 301 TTCATGACGATTGATAAGCTTAAATCCGATCCCTGTAACCCAGCCAGCCAGACAGATTAGAA 360
QY 361 TACCATTATAGAGTCTTAAAGCATAGGCGCATGATCGCAAGTGGGTGGCGAGCTCAA 420
Db 361 TACCATTATAGAGTCTTAAAGCATAGGCGCATGATCGCAAGTGGGTGGCGAGCTCAA 420
QY 421 GTGATGGCAAGTGTCTGCTGAAGCCGAATGGAAGCCATGATTCAGAGAGAGATTAA 480
Db 421 GTGATGGCAAGTGTCTGCTGAAGCCGAATGGAAGCCATGATTCAGAGAGAGATTAA 480

RESULT 5
LOCUS AX788995
DEFINITION Sequence 1459 from Patent WO02066501.
ACCESSION AX788995
VERSION AX788995.1 GI:32955341
KEYWORDS
SOURCE
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE
AUTHORS Legrain,P., Rain,J.C., Colland,F., de Reuse,H. and Labigne,A.
TITLE Protein-protein interactions in Helicobacter pylori
JOURNAL Patent: WO 02066501-A 1459 29-AUG-2002;
Hybrigenics (FR); INSTITUT PASTEUR (FR)
FEATURES
source : l. .853
/organism="Helicobacter pylori"
/mol_type="unassigned DNA"
/db_xref="taxon:210"
ORIGIN
Query Match 61.3%; Score 294; DB 6; Length 853;
Best Local Similarity 100.0%; Pred. No. 7e-67;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
Db 560 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 619
QY 61 CTACCTCACCGCTATCCATGCTTTTAGTGATAGAAATTATAGAGTTACAAGCAATAA 120
Db 620 CTACCTCACCGCTATCCATGCTTTTAGTGATAGAAATTATAGAGTTACAAGCAATAA 679
QY 121 AAAATTGCTTATAGAAATATCACTTTAATGAAGAGCTGTTTAAACGGGCAATTCCT 180
Db 680 AAAATTGCTTATAGAAATATCACTTTAATGAAGAGCTGTTTAAACGGGCAATTCCT 739
QY 181 AATAAGCCCAATTTCCGGCGCTTTGATCGTAGAGGCGATGCGCAACGGGAGGTTT 240
Db 740 AATAAGCCCAATTTCCGGCGCTTTGATCGTAGAGGCGATGCGCAACGGGAGGTTT 799
QY 241 TTAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCGCAAAACAAAAATC 294
Db 800 TTAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCGCAAAACAAAAATC 853

RESULT 6
LOCUS AE017147
DEFINITION Helicobacter hepaticus ATCC 51449 section 4 of 6 of the complete
genome.
ACCESSION AE017147
VERSION AE017147.1 GI:32262489
KEYWORDS
```

```
SOURCE ORGANISM
Helicobacter hepaticus ATCC 51449
Helicobacter hepaticus ATCC 51449
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE
AUTHORS Suerbaum,S., Jogenhans,C., Sterzenbach,T., Drescher,B., Brandt,P.,
Bell,M., Droge,M., Partmann,B., Fischer,H.P., Ge,Z., Horster,A.,
Holland,R., Klein,K., Konig,J., Macko,L., Mendz,G.L., Nyakatura,G.,
Schauer,D.B., Shen,Z., Weber,J., Frosch,M. and Fox,J.G.
TITLE The complete genome sequence of the carcinogenic bacterium
Helicobacter hepaticus
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7901-7906 (2003)
PUBMED 12810954
REFERENCE
AUTHORS Drescher,B. and Suerbaum,S.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Institute of Hygiene and Microbiology,
University of Wuerzburg, Josef Schneider Str. 2, Wuerzburg D-97080,
Germany
FEATURES
source
1. .302212
/organism="Helicobacter hepaticus ATCC 51449"
/mol_type="genomic DNA"
/strain="ATCC 51449"
/db_xref="ATCC:51449"
/db_xref="taxon:235279"
complement (68. .1411)
/gene="ffh"
/locus_tag="HH0940"
complement (68. .1411)
/gene="ffh"
/locus_tag="HH0940"
/codon_start=1
/transl_table=11
/product="signal recognition particle protein"
/protein_id="AAP77537.1"
/db_xref="GI:32262490"
/translations="MFNTLTSPFKGIINKIRFSDDEKSLQACDELKTKLTKNDVYHK
AVKQIIQEVQORTKEAGIKGNFTNALQDALSKILQSKSYGSFSPPTPSVIMMGL
QSGKTSKAKLALYFKQKGVLLVACDLRLAAI EQLTQAQS IVDITFPGHSSQ
PNNAVYAKAKQKATEAHYDIVIIDSAGRLAIDEVLMSELEIKRAVASECLYVAD
SLSGQGIQRAQHPDDKIGIDVILSKFSDTKGGIALSIAYIGIKPFGISGEKIA
DFDIFPDRIISRLMGAGDIASLAKBTANVINBEAKTIARKIKKGQGFDFDIAQIE
NVFKLSSMSIASMIPLGLGMNALKDVIDLKSGEIKNIRAMVNSMTKKRNPDLN
GSRKRAQSGGLEVSIDINRIIKQFQAAKFAKKLSQGGGQELMSMGNLRIPK"
1447. .1557
/locus_tag="HH0941"
1447. .1557
/locus_tag="HH0941"
/notes="predicted membrane-associated"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP77538.1"
/db_xref="GI:32262491"
/translations="MFNRKFYTKITFTMLLPASNIYDYKILYIFPKDYTK"
1554. .1964
/locus_tag="HH0942"
1554. .1964
/locus_tag="HH0942"
/notes="FurR family"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="AAP77539.1"
/db_xref="GI:32262492"
/translations="MMNFEQHLRDKHLKITPQRIATNLQIYNNGHMSIEEIEYQIKQI
YPSISLATIYKVNVALCKANTLREIKAPDKQKQYELSSDKHLHVYCEKGRLLDIKLD
TRALEQNCSSAGSYTISDISAVLMGVCPDCKNAS"
complement (1969. .4014)
/locus_tag="HH0943"
complement (1969. .4014)
/locus_tag="HH0943"
```

```

/note="similar to Helicobacter pylori JHP0417"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77540.1"
/db_xref="GI:32262493"
/translation="MDNAHOTQMOPTSLRAICQSGALHLLSYACRYALNMFNPASI
NPKLFEELKSPAOYVNVYCNCGEDLANIIDVHTNFVYPTKKGKGLKHTMLEONFALSS
NPLNNSATYLHNRALNHLKDEKLPLTAGICNPMIDMLNPEFGNRRHLHSFLPN
MTLFLGGMFQFLIEQSYFHSLLPEHIDFRISLYFVDYPLLFVRSERSCYV
FYKDLHKEFVQYQARRITNNFLRLELCYDSPKQAVQESVAEYAINRSGWGSF
DDEMGKNTFLNLTGNKAHLAPILHLQVRNAPICVVGNGPSLMLLPFTKANAHN
MIFSGTALKEPLKNAGIBPDQIBTERISYLKDVLOSAPLGTLLCGNMVOPDALA
LAEAYIWMRGSASAYFGCAKSVVEPAAPYVGNAGPALACLLSEVLICGLDCGVK
GLSKRAQSFYGEESTIPQNAIVOGNESEVIADALFSLSSAMSKAIGIFPKPV
INLGGAYIRGAPACTDFTLRNIDKSAIILKYSMAHQAVLTRDDCRFQDSI
ITYKEILSLRESVGEISQVRNVAHKVELFRIDIMLYALILKSAREPFVGVLF
EGSISHILTMCLALHPKDDIALFYACQVEIIESGLNMTWSYQMLAMSHNSNL"
complement(4020..5078)
/locus_tag="HH0944"
/complement(4020..5078)
/locus_tag="HH0944"
/note="possible peptidase"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77541.1"
/db_xref="GI:32262494"
/translation="WMRNSSLPYLSLDSAOYVEGFGSCDNALLVCIESRFFITDS
RYTLAEKELCPHTEVIESANFIQSOEVLQKLHIKIVFNPOELSVQMYEKLKNALE
KHKELIPOENPHQKLRICKSEQELGILTKSLKNKAFKRGFYLNKLVKKSPEKE
LHYQAMQPLSHFGEISLFEPIEINGANGAKHALPNKAKYLAKNIDILLFDAGIKYKR
YCSDMTRAAIKGEIHFQKQKPNKYQKIYDVLKAOEEAISKARSGMKAKQIDSL
ARDVTEKSGYGYFVHSTGHGVGLDIHELPRISLSECEVDGMVFSVBPGLYLPQEF
GVRIEDLVVMKNGRAEVL"
complement(5079..5555)
/locus_tag="HH0945"
/complement(5079..5555)
/locus_tag="HH0945"
/codon_start=1
/transl_table=11
/product="3-dehydroquinase dehydratase II"
/protein_id="AAP77542.1"
/db_xref="GI:32262495"
/translation="MKILVIOQPNILNHLRPHFYVGNFTLDQIHNNLQAAQONNAE
LEFFOSNPEGIIDLQECIGEYAGVLINPAAFSHTSIAIADAASCGVPVIEVHIS
NIHAREYRSKSYTCAVAGVITGFGAGYHLALGILQIANEIALAKAQQNNQNA"
complement(5675..7345)
/locus_tag="HH0946"
/complement(5675..7345)
/locus_tag="HH0946"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP77543.1"
/db_xref="GI:32262496"
/translation="MTKKALYEIVDEIGTOSARFGLNI FNLVNFVFSSEYAYCAA
VRATQTPRAYDKHNDKYYGOTFEELDIGQNIKDSIFNTGNTYTTDLADIKKV
RDIRASGKDITKLNTEDEKEDIMANYADEVMKDFKGLNLRKRNNTDTIILDK
PENNVKTAQLKVIKSTNGLKDRYLEGDAADVDELKPPDDYKRHKENLEKMIKEKGN
PENNOVAEKAQAKAEKLANNNCNRLMCENPRTIATVATQSIIVASGHIQAQFSDA
IIAALATLANGVIVWEVKDMFGS IDTETSILKRIKRLKRTIEAFOATFGRGAGFQAI
DAAVGVQIIFRSI TAGKLKLVWDKIRTKLSIYNGIVSVI KGEVSNLRLELIGIILKSL
ESAAWVSTLAEGLKPLILASMAPPLVPFPVFAIILAGAPAVVTHRSIDLVLTLF
GVFAARDRAKURAEIADIIAQKQPALIEDREKLEELTASTHKERLMSLSDSFADYQK
AYQNADDNGIYEGLNRIENLYDSSLISKDMGDVENILKPNRTKGLQ"
complement(7351..8502)
/locus_tag="HH0947"
/complement(7351..8502)
/locus_tag="HH0947"
/note="similar to Helicobacter pylori JHP0703; predicted
membrane-associated"
/codon_start=1

```

```

/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77544.1"
/db_xref="GI:32262497"
/translation="MPLPLIPLILGGAALATAGYGVKXKAI DAYDDSOQTAELYHERAKK
KFTQSGARLDSKSAQVQKRFESLQLOEQIVTDELYADIIVDKLNDADLOIIVG
KEYTQGLVANTQOSIVTSLTGLTGGALAGALTFGMYGVGGLASASTGTAIASLS
GVAATNATWALGGSLSAAGLGMAGTMTVLGGIVAAPIAVMGSSVFAAKAEKKYDA
YAYDSVVAVCETHEAQLVWEQLSKRAKERSTILKESRVVFDASLEKIEDIMENKGY
TDIKNSNDDEHRELMTQLAEVIVAVINAPIMNDEDLTQELIEHQKCKCKLMKRI
QKWQGMESYSEKHEDLMKSFSGGKEYSENDRSIDDDGTCFSGIPQGR"
complement(8561..8836)
/locus_tag="HH0948"
/complement(8561..8836)
/locus_tag="HH0948"
/note="similar to Helicobacter pylori HP0764"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77545.1"
/db_xref="GI:32262498"
/translation="MMINSFPTNCMCKNIGGIKELKSPDLGNGICKYLDQTQSNLC
KIYNRPDICRVDVMPFERYFKYTKBAFYKLNMQSKILQEKERID"
8937..9182
/locus_tag="HH0949"
8937..9182
/locus_tag="HH0949"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP77546.1"
/db_xref="GI:32262499"
/translation="MNKIQDEIDNAYRIIIGLVNKKARELKGLTQLELSLAIGHKSVS
VWSCAIYHKQHFNIHLLKIAQVLEVNQNPFEG"
complement(9329..11815)
/locus_tag="HH0950"
complement(9329..11815)
/locus_tag="HH0950"
/codon_start=1
/transl_table=11
/product="biotin sulfoxide reductase Bisc"
/protein_id="AAP77547.1"
/db_xref="GI:32262500"
/translation="MNVLSRQVLOAMGKLAVPALVPFGCDTESKKAIVESGVKSF
SENVLVDGEVITGAHGVKAKIKDGKVISSTNALKNNVSNPLHSTIPDLIYSPTRIO
YPNVRKSYLENPANRPELGRADWVRVSDAIALAKELKATYKQKKGEGVAGSY
Query Match 34.6%; Score 166; DB 1; Length 302212;
Best Local Similarity 62.2%; Pred. No. 4.4e-33;
Matches 290; Conservative 0; Mismatches 155; Indels 21; Gaps 1;
QY 26 AATCTCAATTTTATAGACATCTTACAAATTCACCTACCGCTATCCCATGCTTT 85
DB 233434 AACCCATAATGAATGGAGAAAATTCGTAATTTTACCACATCGTTATCGATGTTAC 233493
QY 86 TAGTGTAGAAATATAGAGTTACAACCAATAAAA-----AAA 124
DB 233494 TTGTTGATAGATTAGGAGCTTGTGCAAAATACAAGTATAGTGAACCAATGTTATA 233553
QY 125 TTGTCGCTTATAGAATATACATCTTTTAATGAAGACGTTGTTTAAACGGGCATTTCCCTAATA 184
DB 233554 TCAAAACATATAAATATGAACAAATTAATGAGAGGATTTTTCGGGCATTTTCCAAACA 233613
QY 185 AGCCCATTTTCCGGCGGTTTTCATCGTAGAGGCATGGCGCAACACGGAGGGTTTTTAG 244
DB 233614 AACCTATTATCCGGGTGAATGCAGATTGAAGGAATGGCGCAACGGCGGTCCTTTAG 233673
QY 245 CCTTCACTAGCTGTGGGGTTTGGACCTGAATTCGCCCAACCAAAATCGGTATTTCA 304
DB 233674 CATTTGTAAGTATGTTTGGTGAATAATGTTTCGAGGAGGCAAAAATAAATTTGATATTTTA 233733

```



```

Oy 305 TGACGATTGATAAGTTAAATTCCTCGCATCCCTGTGTAACCCAGCGACAGATTAGAAATACC 364
Db 233734 TGACCATAGATAAATGTTAAAGTTTCCTGTATCCCTGTCTGCCAGGTGATAGCTTGTATATG 233793
Oy 365 ATTAGAGAGTCTTAAGCATAGAGGCATGATCTGCGCAAGTGGTGCACGGCTCAAGTGG 424
Db 233794 AGCTCAAGGTTTAAACACAAAGGAAGTATATGGAACCTTGGAGCAATGCTTTTGTG 233853
Oy 425 ATGGCAAAAGTGTCTCCTGAAGCCGAATTGAAAGCCATGATTTCGAGA 470
Db 233854 AGGATAAACTTGTAAGTGAAGCTGAACCTTAAAGCTATGATTACTGA 233899

RESULT 7
BX571657/c 347961 bp DNA linear BCT 17-APR-2005
LOCUS Wolinella succinogenes, complete genome; segment 1/7.
DEFINITION BX571657 BX571656
ACCESSION BX571657.1 GI:34482172
VERSION complete genome.
KEYWORDS Wolinella succinogenes
SOURCE Wolinella succinogenes
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Wolinella.
REFERENCE
1 Baar,C., Eppinger,M., Raddatz,G., Simon,J., Lanz,C., Klimmek,O.,
Nandakumar,R., Gross,R., Rosinus,A., Keller,H., Jagtap,P.,
Linke,B., Meyer,F., Lederer,H. and Schuster,S.C.
Complete genome sequence and analysis of Wolinella succinogenes
Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11690-11695 (2003)
14500908
2 Schuster,S.C.
Direct Submission
Submitted (15-MAY-2003) Max-Planck Institut for Developmental
Biology, Spemannstr. 35, 72076 Tuebingen, GERMANY
Location/Qualifiers
1. 347961
/organism="Wolinella succinogenes"
/mol_type="genomic DNA"
/strain="DSMZ 1740"
/db_xref="taxon:844"
1. .1314
/gene="dnaA"
/locus_tag="WS0000"
1. .1314
/gene="dnaA"
/locus_tag="WS0000"
/locus_start=1
/codon_start=1
/transl_table=11
/product="dnaA, probable chromosomal replication initiator
protein"
/protein_id="CAE09174.1"
/db_xref="GI:34482173"
/db_xref="GOA:Q7MSY2"
/db_xref="InterPro:IPR001957"
/db_xref="UniProt/Swiss-Prot:Q7MSY2"
translation="MLGDTTLKQLKDEISTLEYRYIKOLRYDEEASRTDLAVIAPN
QPIANWVKTYGERLAHLFELSTGIRPKIEIRLGLSKDVKSSSPKAGVSKGQKSTIL
NPSFTPDSFVGVGNRFAVEVSONVAKQCIAYNPILLIYGGTGLGHLINSIGNYV
AKGKSVIYVTSQFLNDLYLHIRNNTMDRFRDKYACDYLIDVQFFGKPKQIQEAF
FHTFELHNKNQIQLTSDKTPQIAGLEELKSRFEMGVSDIQPELETKNIIKX
KECFDGLUSNEIISYATNMNNINREIGIILKNAYANLMNQEITLQPAKNVLKQ
IKKEENITLENIIEVVSKELVNPKPAEIKSGSKSNANRIVIFLARTLTNPSMPS
LAQFFGMDHSSVSKAMTKVTKEEDANFQWVIEELKNKIKSRN"
1474. .2541
/gene="dNAN"
/locus_tag="WS0001"
1474. .2541
/gene="dNAN"
/locus_tag="WS0001"
/EC_number="2.7.7"
/codon_start=1

```

```

/transl_table=11
/product="DNA POLYMERASE III, BETA CHAIN"
/protein_id="CAE09175.1"
/db_xref="GI:34482174"
/db_xref="GOA:Q7MAV2"
/db_xref="InterPro:IPR001001"
/db_xref="UniProt/TREMBL:Q7MAV2"
translation="MNIVIPKSSLENILTNLPFLFKKASQITSHVYFKAQGSSELPF
QATFYIGLLAKTESVQIQEGIATVNGKKILDIVRLEKEGITLQOEYLIHKQEK
SFKLPFMNAEEPPSPSEELPKIEINSMHLVQSMKKILPAIDNNPKFELNGALLD
IKESYFNVAATDTRRLAVIKYENPINKLALIVPKALVRIQKLFFEEVLEHSETHL
ILKSNYLFPTKLISGKYDPYERLIIPKETHNMTLPKEKLEIKLINSLSNDIKVTF
TKEEILFESLSSENSEAKTQVEVTGITKSLTGLGINSKYLDDYLGQIDTSEFKIGINE
SNTPTFLKSENFSTVIMPVM"
2561. .4879
/gene="gyrB"
/locus_tag="WS0002"
2561. .4879
/gene="gyrB"
/locus_tag="WS0002"
/EC_number="5.99.1.3"
/codon_start=1
/transl_table=11
/product="DNA GYRASE SUBUNIT B"
/protein_id="CAE09176.1"
/db_xref="GI:34482175"
/db_xref="GOA:Q7MAV1"
/db_xref="InterPro:IPR001241"
/db_xref="InterPro:IPR002288"
/db_xref="InterPro:IPR003594"
/db_xref="InterPro:IPR006171"
/db_xref="InterPro:IPR011557"
/db_xref="InterPro:IPR011558"
/db_xref="UniProt/TREMBL:Q7MAV1"
translation="MQNQYSDANIKVLKGLVAVRKEPMYIGDTNVNGLHLIYEVV
DNSIDEAMAGHCNTISITLTKESAIIEDNGRGI PVDMPHTENI PAATVLTVLHAGG
KFDKSYKSGGLHGVSVNALSCLKLVMTRKNGNVYTFQFSGCPKTDGLDILGTS
KNGTTFIETIPDDEIFEVTVFDRFLARRFKELAYLNRTIHPKDERDNFESYHFE
GGLRQFVHLNKKVLVSQVFSFSDNESDTEVEIALYNDGYDEKVLSPVNNIKTPEGG
THEAGFRAGLSRAITSVIDANANAREKDNKITGDDVREGLVAIVSVKMEPQEGQTK
GKLGSFVRPIVQKLTYEKIAKFEENPNDAKAIMQKALMAAREAKAKARELTKK
DSLMVTLPKLDACQKPSISELYLVEGDSAGSAGQDRDFVQALPLPKGLINV
EKSRLDKLSDEIKNMITALGCGIEEFDIRIRYQKIIMTDADVDGSHIQTLILT
FPRYLKPVIEAGLYIAQPLYRFKKGKEIYLRDEKALSEYLIENGIFAFEGVG
TPDLLEIRPASHVRLTLEKELKPFAMIDVVKVMIEHRDLISLPMSELYIKLEBEVQQ
KGYNILNKVEEEKIHLVQTHEGLVDLVDLSTFTHLPFEAYVHNLIERDLSEFL
GERDILTLEIEESAKGAYIQRYKGLGEMNPEQLWETTPENRLLKVKQITDDME
ASDVFSLFMGDEVEPRRAYIQDHAKDVKHLDV"
4890. .5264
/locus_tag="WS0003"
4890. .5264
/locus_tag="WS0003"
/notes="HYPOTHEITICAL PROTEIN CUI1724C"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAE09177.1"
/db_xref="GI:34482176"
/db_xref="GOA:Q7MSY1"
/db_xref="InterPro:IPR001474"
/db_xref="UniProt/TREMBL:Q7MSY1"
translation="MRYGEMVKNFPEEIEVWPNRNDRHVYIKITLPEFSCLCPRSG
YPIYATVYIEYVSSLAVELKAIKLYINSFDRHVSHEDSANEIYDLLYKKLSPKELY
LRMDFNPRGNVHTIIIEIDSKKN"
5274. .5900
/locus_tag="WS0004"
5274. .5900
/locus_tag="WS0004"
/codon_start=1
/transl_table=11
/product="PUTATIVE INTEGRAL MEMBRANE PROTEIN"
/protein_id="CAE09178.1"
/db_xref="GI:34482177"

```

```
/db_xref="GOA:Q7MAV0"
/db_xref="InterPro:IPR003811"
/db_xref="UniProt/Swiss-Prot:Q7MAV0"
/translation="MSFISNPNIHFYLLAYFIGIPFGYLLSRFFYNINIOEVGSGSI
GATNVLRLKEDPKAKMMWALTMLLDALKALVILVAKMGMSVETOWAIAVLAVV
GHCFPYLKEGKGKGVATGLGVNMAVLPLEAAIGLVVVLVVGKTKLSSLSLLGLGA
LISPSLIHPIEPIGTHAPLLIIAFVIFYKHIPNLVRLFGQKEGRVV"
5897..6220
/locus_tag="WS0005"
CDS
5897..6220
/locus_tag="WS0005"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CA809179.1"
/db_xref="GI:34482178"
/db_xref="GOA:Q7MSY0"
/db_xref="InterPro:IPR006157"
/db_xref="UniProt/TREMBL:Q7MSY0"
/translation="MTLLEHLKEALIGILPEREKAQIILVDGSPSYVPAGNYLDY
VAIKELITQHLISERFGLLEALVSLHRELKESFEELTEVNLTLKRPDILLRDLVGAS
LSKHY"
complement(6224..8179)
/locus_tag="WS0006"
complement(6224..8179)
/locus_tag="WS0006"
/note="PROBABLE CHOLINE TRANSPORTER"
/codon_start=1
/transl_table=11
/product="PUTATIVE CHOLINE TRANSPORTER"
/protein_id="CA809180.1"
/db_xref="GI:34482179"
/db_xref="GOA:Q7MAU9"
/db_xref="InterPro:IPR000060"
/db_xref="UniProt/TREMBL:Q7MAU9"
/translation="MKIRATILPPVFSSILLIFGVFLGALMPQANTILFSLMIQKGI
IKELGWFIYSVALDFIVLITLAISSRDGVKLGADHSEPEFSYKSFAMLFSGMGIG
LMFYGAEPVHFLSPKKEGTVAAEKAMKTFPHFGIHWAIYALVALALAYFSF
RHLQFTAPSPYFPIGIRIKRGLADITFAVITMFGVSLGFGVLQVNSGLSYL
FIPIWTOIILLAFITSLATISVVLGSGGKRLSELMLFLAIALMLFVLLAGPTL
FIQTPVONIGLAVLDTYKTNLHAYEPTGIGWTIFVMAWIAWSPVGMFIARI
SKGRIREFVIGLVMPVGFTELMLTVFGDSALFILKENLTLGSAVQADSTALPK
FLESPPFITSLLATQITLITFFVSSDGLVDMLTGGIKETIPMQRIFWSSSEG
FYAALLAGAGLQATISIFALPFSIIMLFMAYGMHRLRIEAIKKEALKVQGAIT
FSPSPYPMQWRLKILSRHKSAITLQIRVFEAITAVSOELREGLVSKVKGGE
GSIISLVHGREIDFLYERNRGYTTTFIMSESEDEIYQYHRAEVFLKEGSGQYDL
MGYSKEQIADILNQYRHWQFLHLLR"
complement(8279..9136)
/locus_tag="WS0007"
complement(8279..9136)
/locus_tag="WS0007"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CA809181.1"
/db_xref="GI:34482180"
/db_xref="GOA:Q7MSX9"
/db_xref="InterPro:IPR006015"
/db_xref="InterPro:IPR006016"
/db_xref="UniProt/TREMBL:Q7MSX9"
/translation="MLSSKFPILLVATDLSPTGKTVLEKAFKLAKRYHKPLFIHVSFT
TLFAPKIELTQIEANLCKLEPFFVARYPEVKATILCROGGIEESIISAIKETOASLL
LIGTSGENILAREFFLGSHCRKIIRSTTIPLVVKNKEESDYSKIPIPTDPSPOKM
ANLTLHPPHAYISLPHVLKPFQRLGMYGLDKDBILRYHKDIDNQSEADHFIEK
LKPSPGKRILKMLMETGVFSAKLQKVKAEQSDLIALLNTGNISFFSMDILNQSPKD
VLIFKTDSE"
complement(9299..9790)
/locus_tag="WS0008"
/locus_tag="WS0008"
Query Match 34.4%; Score 165.2; DB 1; Length 347961;
Best Local Similarity 61.3%; Pred. No. 7.1e-33;
Matches 266; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
```

```
Qy 47 ATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTAGTGATAGATATATAGAGT 106
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62061 AGATTAAAGAGATCTTCCCATCGAATCCCTTTCTTATTTGGTTGATCGAGTTACGGCAC 62002
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 TACAAGCCCAATAAAAAATGTCGCTTATAAGAAATATCACTTTTAAATGAAGAGCTGTTTA 166
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62001 TCACCTCGAAGAGAGCATCGAGGCTACAAAATATACCAATCAACGAAGAGTGTTC 61942
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 ACGGCGATTTCCCTAATAAGCCCAATTTTCCCGGGCGTTTGTATCTGTAGAGGGCATGGCGC 226
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61941 AAGGCCATCTCCCATCAAGCTGTCTATCTGCGGTACTCGTCAATTGAAGGTATGGCTC 61882
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 AAACGGGAGGGTTTTTAGCCTTCACTAGCTTGTGGGGGTTTGACCTTGAAATCGCCAAA 286
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61881 AAGCAGGAGGCGTGTGCTTTTGTAAAGTATGTTTTGGCGAAGAGCGCAATCACGATG 61822
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
287 CAAAAATCGCTATTTCATGACGATTGATAGTTAAATTCGCGATCCCTGTAAACCCCGAG 346
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61821 AAAAAATCGTCTACTTCAATGATATTGACAGCAAGTTTCGCGTCCCGTCAACCCCGAG 61762
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
347 GCGACAGATTAGAATACCAATTTAGAAATCTTAAAGCATAAAGGCGCATGCTGGCAAGTGG 406
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61761 GAGATAAATCGTCTATCGCTCAATGCTTAAGCACAAAGGAGCAATTGGATTCTCG 61702
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
407 GTGGCAGCGCTCAAGTGGATGCGCAAGTGGTCTGCTGAAGCCGAAATTGAAAGCCCATGTTG 466
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61701 AAGCAGCGCTTATGTTGATGACAAACTCGTCCGCAAGCGCAACTCAAAGCCATGGTGC 61642
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
467 CAGAGAGAGATTAA 480
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61641 CTGCAAGGAAAAA 61628

RESULT 8
CJ11168X1
LOCUS Campylobacter jejuni subsp. jejuni NCTC 11168 complete genome;
DEFINITION segment 1/6.
ACCESSION AL139074 AL111168
VERSION AL139074.2 GI:6967505
KEYWORDS
SOURCE Campylobacter jejuni subsp. jejuni NCTC 11168
ORGANISM Campylobacter jejuni subsp. jejuni NCTC 11168
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 Parkhill,J., Wren,B.W., Mungall,K., Ketley,J.M., Churcher,C.,
Basham,D., Chillingworth,T., Davies,R.M., Feltwell,T., Holtroyd,S.,
Jagels,K., Karlyshev,A.V., Moule,S., Pallen,M.J., Penn,C.W.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., van Vliet,A.H.,
Whitehead,S. and Barrell,B.G.
The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences
Nature 403 (6770), 665-668 (2000)
10688204
2 (bases 1 to 314150)
Parkhill,J.
Direct Submission
Submitted (09-FEB-2000) Submitted on behalf of the Campylobacter
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
Notes:
Details of C. jejuni sequencing at the Sanger Centre are available
on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/c_jejuni/).

FEATURES
source
1. 314150
/organism="Campylobacter jejuni subsp. jejuni NCTC 11168"
/mol_type="genomic DNA"
/strain="NCTC 11168"
/db_xref="taxon:192222"
1. .1323
gene
```



```

/misc_feature
1. .1323
/gene="dnaA"
/gene="Cj0001, dnaA, probable chromosomal replication initiator protein, len: 440 aa; similar to many e.g. DNAA_ECOLI (467 aa), fasta scores; opt: 839 z-score: 949.6 E(): 0, 32.3% identity in 470 aa overlap. 39.9% identity to Hp1529. Contains P50017 ATP/GTP-binding site motif A (P-loop) and Pfam match to entry PF00308 bac_dnaA, Bacterial dnaA protein, score 419.90, E-value 4e-141"
/codon_start=1
/transl_table=11
/product="chromosomal replication initiator protein"
/protein_id="CAB72494.1"
/db_xref="GI:6967506"
/db_xref="GOA:Q9PJ00"
/db_xref="InterPro:IPR001957"
/db_xref="InterPro:IPR003593"
/db_xref="InterPro:IPR010921"
/db_xref="UniProt/Swiss-Prot:Q9PJ00"
/translacion="WNPSQILENLKKESENEYENLSLKFNEKQSKADLLVFNAPN ELMAKFIOTYKGYKTAHFYVQSGNKAIINIOAQSAQSNKSTKIDIAHKAQSTILN PSFTESFVGDGSKYVACAKAIAHKDKLGLNPIFYVGTGLKTHLLQAVGNAS LEMGKKVIYATSENFINPTSCNLSKGLDKFHEKVRNCVDLLIDVQFLGDKTKIQEE FFFIFNEIKNDGQIMTSDNPPNMLKGIETELKGRFAHGIADITPPOLDTKIALIR KKCFENDLSNDIINYIATSKDNIETIEGIIISLNAYATILGQETITLELAKSVMKD HIKEKENITIDILSLVCKEFNIXPSVSKNKTQNIIVTARRIYILARALTALTMP QLANFYEMKDHTAISHNVKKITEMIENDASLAKAKIEELKNKILVKQSQ"
301. .1239
/gene="dnaA"
/notes="Pfam match to entry PF00308 bac_dnaA, Bacterial dnaA protein, score 419.90, E-value 4e-141"
427. .450
/gene="dnaA"
/notes="P50017 ATP/GTP-binding site motif A (P-loop)"
1483. .2550
/gene="dnaA"
1483. .2550
/gene="dnaA"
/EC_number="2.7.7.7"
/notes="Cj0002, dnaN, probable DNA polymerase III, beta chain, len: 355 aa; similar to e.g. DP3B_ECOLI DNA polymerase III, beta chain (EC 2.7.7.7) (366 aa), fasta scores; opt: 349 z-score: 372.3 E(): 1.7e-13, 19.6% identity in 367 aa overlap. 34.2% identity to Hp0500. Contains Pfam match to entry PF00712 DNA_pol3 beta, DNA polymerase III beta subunit, score 119.80, E-value 5.2e-32"
/codon_start=1
/transl_table=11
/product="DNA polymerase III, beta chain"
/protein_id="CAB72495.1"
/db_xref="GI:6967507"
/db_xref="GOA:Q9PJA9"
/db_xref="InterPro:IPR001001"
/db_xref="UniProt/TrEMBL:Q9PJA9"
/translacion="MKLSINKNTLESAVILCNAYVVEKDSSTTJSHLFFHADEKLLI KASDVEIGINYKIKRIVSSGPTANAKSIADVIKSLNVEVLTIDNPLFVRQKS TKYKLPFMRHEDFPNPTENGNQFDISDLSRLKPLPSIDTNPKYSLNGAFID IKTDKINFVTGDTKLRAIYTLKANNQFDSFSIPKAIEMQKLFYEKIEIFYDNML IAKNFNFEFTKLINDFPDYKVIPTKPKGLSFSTEDFIDSLKISVVTXKRLHF NKDKIIIFEGISLDNMEAKTELEITQGVSEFNLTIKIKLLDLFLTSIEBEKFTLSVNE PNSAFIVKSGLSLMIIMPIL"
1483. .2535
/gene="dnaN"
/notes="Pfam match to entry PF00712 DNA_pol3 beta, DNA polymerase III beta subunit, score 119.80, E-value 5.2e-32"
2579. .4888
/gene="gyrB"
2579. .4888
/gene="gyrB"
/EC_number="5.99.1.3"

```

```

/notes="Cj0003, gyrB, probable DNA gyrase subunit B, len: 769 aa; similar to many e.g. GYRB_EACSU DNA gyrase subunit B (EC 5.99.1.3) (638 aa), fasta scores; opt: 2130 z-score: 3165.8 E(): 0, 47.7% identity in 773 aa overlap. 61.2% identity to Hp0501. Contains P500177 DNA topoisomerase II signature and Pfam matches to entry PF00204 DNA_topoisolI, DNA topoisomerase II (N-terminal region), score 915.20, E-value 1.9e-271, and to entry PF00986 DNA_gyraseB_C, DNA gyrase B subunit, carboxyl terminus, score 154.00, E-value 2.6e-42"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="CAB72496.1"
/db_xref="GI:6967508"
/db_xref="GOA:O87667"
/db_xref="InterPro:IPR000565"
/db_xref="InterPro:IPR001241"
/db_xref="InterPro:IPR002288"
/db_xref="InterPro:IPR003594"
/db_xref="InterPro:IPR006171"
/db_xref="UniProt/Swiss-Prot:O87667"
/translacion="MOENYCASNIKVLKGLKLEAVRKPQWYIGDTNIGLHMHVYEVVD NSIDEAMAGHCDTIDVEITTEGSCIVSDNGRGI PVDMPHTENMPTLTVLTVLHAGGK PKDITYKVGSLGCVGVSVNALSKKLVATVERNGEIQRFESGKVIISFEGVIGKSK KTGTTIEFWPDDQIFEVTEFDYELAKRFELAYLNPKITINFKNRVRGKHESHFEG GISQFTDLNKEALTKAIFFSVDEEDVNVVALLYNDTVTSNLLSFVNNIKTPDGGT HGAQFMGLTRVTSNIYEAASAKDNKIKTDGDRGLIAIVSVKPEPQEGQTKG KLGSTVVRIPSKASFEYLTYFEENPIEAKIMKALMAARGEAKKARELTKRKE SILSVGLPKGLADQCSKDPSESEIYLVGDSAGSAGKAGRSFQAILPLRGKILNVE KARLDKILKSEQIQNMITARGCGIGEDFDLSKLRHKIIMTDADVDGSHQITLLTF FRFMNVLVANGHIYLAQPPLYKKAKQIYLKDEKALSEYLLTIEGLNLYEGIGM NDLKQIKIYAAYRAILKDLKGFNVISVYMIENSLVKNNEELFSVKQLETO GHNLIHYNENEIRAFVQONGLELVINEELFTHPLYEEASYIFDKIKDRSLFEDK DILVLEDVETNAKKGATIQRYKGLGEMNPEQLWETTMDSVRLLLKTIIEDAQSAND TFNLFMGDEVEPRDRIQAHAKVDKHLV"
2606. .4168
/gene="gyrB"
/notes="Pfam match to entry PF00204 DNA_topoisolI, DNA topoisomerase II (N-terminal region), score 915.20, E-value 1.9e-271"
3830. .3856
/gene="gyrB"
/notes="P500177 DNA topoisomerase II signature"
4649. .4852
/gene="gyrB"
/notes="Pfam match to entry PF00986 DNA_gyraseB_C, DNA gyrase B subunit, carboxyl terminus, score 154.00, E-value 2.6e-42"
complement(4916. .5257)
/gene="Cj0004c"
complement(4916. .5257)
/gene="Cj0004c"
/notes="Cj0004c, possible periplasmic protein, len: 113 aa. No Hp match. Contains possible N-terminal signal sequence"
/codon_start=1
/transl_table=11
/product="putative periplasmic protein"
/protein_id="CAB72497.1"
/db_xref="GI:6967509"
/db_xref="GOA:Q9PJA7"
/db_xref="UniProt/TrEMBL:Q9PJA7"
/translacion="MKKILILALFLSASQAQNLINPDITGLIIDPDSPLVANCALAC HGSNLIYNMHASRKAWLAARWMQDSEGLWIEPEDEKILNLEYKGYKDYTRRRI FLAILQNKTH"
complement(5260. .6498)
/gene="Cj0005c"
complement(5260. .6498)
/gene="Cj0005c"
/notes="Cj0005c, possible molybdenum containing oxidoreductase, len: 412 aa; similar to many eukaryotic oxidoreductases e.g. SUOX_HUMAN sulfite oxidase precursor (488 aa), fasta scores; opt: 543 z-score: 338.2 E():

```

```
1.4e-11, 29.8% identity in 372 aa overlap and NIA_PETHY
nitrate reductase from Petunia hybrida (909 aa), fasta
scores; opt: 281 z-score: 312.5 E(): 3.8e-10, 28.9%
identity in 395 aa overlap. No Hp match. Contains Pfam
match to entry PF00174 oxidoredo molyb, Oxidoreductase
molybdopterin binding domain, score 45.20, E-value 1e-12"
/codon_start=1
/transl_table=11
/products="putative molybdenum containing oxidoreductase"
/protein_id="CAB72498.1"
/db_xref="GI:6967510"
/db_xref="GOA:Q9PUA6"
/db_xref="InterPro:IPR000572"
/db_xref="InterPro:IPR000566"
/db_xref="InterPro:IPR008335"
/db_xref="UniProt/TrEMBL:Q9PUA6"
/translation="MKQNDQKENRRDFLKNIIGLFLGISVLNFSNPFSLGSKALAKE

Query Match      28.2%; Score 135.2; DB 1; Length 314150;
Best Local Similarity 58.4%; Pred. No. 6.3e-25;
Matches 258; Conservative 0; Mismatches 178; Indels 6; Gaps 1;

QY 39 TATAGAGCATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTACTGGATAGAAAT 98
DB 250598 TGTAAATGCAAAATCAAGAAATCTTACACACCGCTTATCTTTTACTAGTAGATAAAAT 250657

QY 99 TATAGAGTTACAGCCCAATAAAAAATTCGCTTATAAGAAATATCATCTTTTAATGAAGA 158
DB 250658 TACAGAAATTAAGATTAAAGAGTGGTCTAGGATATAAAACATCAGTATAGTGACCA 250717

QY 159 CGTGTTTAAGGGGCAATTCCTTAATAAGCCCATTTCCCGGGCGTTTGTATCGTAGAGGG 218
DB 250718 TGTGTTTTATGGGACATTTTCCAGGCGATCTATTTATCTCGAGTGTGTGATTTTGAAGG 250777

QY 219 CATGGCGCAACGGGAGGTTTGTAGCTTCACTAGCTTGTGGGGTTTGACCTGAAAT 278
DB 250778 TATGGCTCAACAGGTGGAGTTTGTAGCTTTTGAAGTAT-----GGAAGATAAGTGGA 250831

QY 279 CGCCAAAACAAAAATCGTGTATTTCATGACGATTGATAAGTTTAAATTCGCGATCCCTGT 338
DB 250832 TCCAAAAGTAAAGTAGTTTATTTACAGGCATAGATGAGCAAAATTTAGAATCTGT 250891

QY 339 AACCCAGCGGACAGATTAGAAATACATTTAGAAGTCTTTAAAGCATAAGGCGCATGATCG 398
DB 250892 GCGTCTCGAGATAGGCTTGATTTGAAATGAGCGTGGTTAAAAATCGTGTAAATATGTG 250951

QY 399 GCAAGTGGTGCGACCGCTCAAGTGAAGTGGCAAGTGGTGGTGAAGCCGAATTGAAGC 458
DB 250952 GATTTTAAAGGGCAAGCTTTTGTAGATGGAATTTAGTTGCAGAGCGCGAGCTTAAAGC 251011

QY 459 CATGATTGCAGAGAGAGATTAA 480
DB 251012 TATGATAGTGGATATAAATGA 251033
```

RESULT 9  
CP000025\_02  
WFCOMENT

Sequence split into 18 fragments LOCUS CP000025 Accession CP000025

Fragment Name	Begin	End
CP000025_00	1	110000
CP000025_01	100001	210000
CP000025_02	200001	310000
CP000025_03	300001	410000
CP000025_04	400001	510000
CP000025_05	500001	610000
CP000025_06	600001	710000
CP000025_07	700001	810000
CP000025_08	800001	910000
CP000025_09	900001	1010000
CP000025_10	1000001	1110000
CP000025_11	1100001	1210000
CP000025_12	1200001	1310000

CP000025\_13 1300001 1410000  
CP000025\_14 1400001 1510000  
CP000025\_15 1500001 1610000  
CP000025\_16 1600001 1710000  
CP000025\_17 1700001 177831  
Continuation (3 of 18) of CP000025 from base 200001 (CP000025 Campylobacter jejuni RM1221)

Query Match 27.8%; Score 133.6; DB 1; Length 110000;

Best Local Similarity 58.1%; Pred. No. 1.8e-24;  
Matches 257; Conservative 0; Mismatches 179; Indels 6; Gaps 1;

QY 39 TATAGAGCATATCTTACAAATCTTACCTCACCGCTATCCCATGCTTTTACTGGATAGAAAT 98  
DB 86911 TGTAAATGCAAAATCAAGAAATCTTACACACCGCTTATCTTTTACTGGTATAGATAAAAT 86970  
QY 99 TATAGAGTTACAGCCCAATAAAAAATTCGCTTATAAGAAATATCATCTTTTAATGAAGA 158  
DB 86971 TACAGAAATTAAGATTAAAGAGTGGTCTAGGATATAAAACATCAGTATAAGTGATCA 87030  
QY 159 CGTGTTTAAGCGGCAATTCCTTAATAAGCCCATTTTCCGGGCGTTTGTATCGTAGAGGG 218  
DB 87031 TGTGTTTTATGGGACATTTTCCAGGCGATCTTATTTATCTCGAGTGTGTGATTTTGAAGG 87090  
QY 219 CATGGCGCAACGGGAGGTTTGTAGCTTCACTAGCTTGTGGGGTTTGACCTGAAAT 278  
DB 87091 TATGGCTCAACAGGTGGAGTTTGTAGCTTTTGAAGTAT-----GGAAGATAAAGTGGA 87144  
QY 279 CGCCAAAACAAAAATCGTGTATTTCATGACGATTGATAAGTTTAAATTCGCGATCCCTGT 338  
DB 87145 TCCAAAAGTAAAGTAGTTTATTTACAGGCATAGATGAGCAAAATTTAGAATCTGT 87204  
QY 339 AACCCAGCGGACAGATTAGAAATACATTTAGAAGTCTTTAAAGCATAAGGCGCATGATCG 398  
DB 87205 GCGTCTCGAGATAGGCTTGATTTGAAATCAGCGTGGTTAAAAATCTGGTGAATATGTG 87264  
QY 399 GCAAGTGGTGCGACCGCTCAAGTGAAGTGGCAAGTGGTGGTGAAGCCGAATTGAAGC 458  
DB 87265 GATTTTAAAGGGCAAGCTTTTGTAGATGGAATTTAGTTGCAGAGCGCGAGCTTAAAGC 87324  
QY 459 CATGATTGCAGAGAGAGATTAA 480  
DB 87325 TATGATAGTGGATATAAATGA 87346

RESULT 10

AE015936

LOCUS Clostridium tetani E88, section 1 of 10 of the complete genome.

DEFINITION Clostridium tetani E88, section 1 of 10 of the complete genome.

ACCESSION AE015936

VERSION AE015936.1

KEYWORDS GI:28202269

SOURCE

ORGANISM

Clostridium tetani E88

Clostridium tetani E88

Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

Clostridium.

REFERENCE 1 (bases 1 to 300029)

AUTHORS Bruggemann,H., Baumer,S., Fricke,W.F., Wierze,A., Liesegang,H.,

Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and

Gottschalk,G.

The genome sequence of Clostridium tetani, the causative agent of

tetanus disease

Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)

REFERENCE 2 (bases 1 to 300029)

AUTHORS Bruggemann,H., Baumer,S., Fricke,W.F., Wierze,A., Liesegang,H.,

Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and

Gottschalk,G.

Direct Submission

Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute

of Microbiology and Genetics, Georg-August University,

Grisebachstr. 8, Goettingen 37077, Germany

FEATURES

Location/Qualifiers

source 1..300029

```
/organism="Clostridium tetani B88"
/mol_type="genomic DNA"
/strain="Massachusetts"
/sub_strain="B88"
/db_xref="taxon:212717"
complement(1..2304)
/gene="mecB"
/locus_tag="CTC00060"
complement(1..2304)
/gene="mecB"
/locus_tag="CTC00060"
/codon_start=1
/transl_table=11
/product="negative regulator of genetic competence mecB"
/protein_id="AAO34716.1"
/db_xref="GI:28202270"
/translacion="MKLCSVKNKVATLFTTFETGKPOLGICAECAKXNNIOPLNO
IIEAGVTPDEMEYAEQNKLLDNMDINELLNNDLFSNMNNMNGMMFKVQDDFID
KDKNEENTNSRNDKSGNSNVKTRVPKTKRRLDKYAINLNDKADKIDIVVGR
DKETIRIITLNRKNPNPILIGEPGVGKTAEGIAVKIIEKNVPYIKLFDKBIYLLD
LTSVAGTQFRQFGRMAKIVDEVTESENVILVIDEHNITICAGEAQCAGALNAAIL
KPALARGIOVIGATLLEKRYIEKDSALERRFQILVEBPSVKETIILNNIKIY
EDYHKVTKSKLIEBAVKLSRYITDRYLPDKAIDVIDEASSRVNLRNKNLVEFKIAK
BELGLQNMALNEANSNGYKRAAQYKTEYKLNKRVBELEKSKVELSVDDIAYVIEA
WTKIPVQVRKEAEKLNLEQLRHLKRLIGQDKAVRKLRSITRRNRLGFKSKRRPASP
IFVGTGVGKTELVSRLACELFGDEDSLIRIDMSYMEKHTASKLIGSPGVGYDEG
QLTEKIRKPYSVLLDEIEKAHPDIFNMLQVLEDGRDLTDNQGRTVFFDNVSIIMT
SNLNTNFKESSIGFNTWSIQIENSIKDALKETFKPEFINRIDEIVIPEKLNKEELYD
IIDIMFVRSDELSEKNIKVHFTKEVKQPIKLVGVYDEKYGABELKRAIQRYIEDEISE
QYFRNEVREGTNILVDIYEESIYIKNM"
2487..3923
/locus_tag="CTC00061"
2487..3923
/locus_tag="CTC00061"
/EC_number="2.7.8.-"
/codon_start=1
/transl_table=11
/product="putative cardiolipin synthetase 2"
/protein_id="AAO34717.1"
/db_xref="GI:28202271"
/translacion="WKISFTSLSPFFFINIIITLVIILERKRREKTIAMLIITLL
PPGILFLVFLGRNWKHLNNDPSPKIRELIYRSMKIDDDPYVPLIHLIAENSESP
LFLNNSIFDKTEGFKALKEKRLNAKHIIHLEYIVKNDIIGNEIKNLLIEKSLG
VEVRPILDRVGSIKLGEPKINDIMKNGIDVQVSYFLAPLLKKNITQVNNYRHRKLI
IDGVEFTGGINIGDEVYLGKGLGYWRDTHIMVKGDFVLGLOAFYDDFWTIKXAND
SILINTDHSYFPKYNQSGKKLMQLAKSGPSPYSIMQGLIKMLYMA2DHIYITPY
FVPSSEIMALKIACLGIDVRILFPQGDHFTVTSASRYIAELLSKSAKVHFYDEN
SFVHAKIITVQKVTGTGTANNDIRSFDLNIEINALIYDEETIITKLEEMFFDLNHSK
NVTQDFNETSNFMKFIEAITRLFSLL"
4027..4548
/locus_tag="CTC00062"
4027..4548
/locus_tag="CTC00062"
/codon_start=1
/transl_table=11
/product="DNA polymerase, bacteriophage-type"
/protein_id="AAO34718.1"
/db_xref="GI:28202272"
/translacion="WMSGGIPCDILFIGBAPKGTVEEGKPFVGAAGKNFQYVLDMIN
LNKDIRITNTCYFRPFIKMSSSGRKTSNRPPTKEINLFQHLKKEIETVNPXII
ITLGNTPKTKTQVNNATIGHRHGSLLFSEKLNHIFPMVHPASALIYNRTDSFKESYKE
DWCKLADILKGPK"
complement(4584..5144)
/locus_tag="CTC00064"
complement(4584..5144)
/locus_tag="CTC00064"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO34719.1"
/db_xref="GI:28202273"
/translacion="MYDITYTIERGINMSIKNSIGLSKIATKTVGDGANTVAQKSS
EFVEISKLNLSSEKNNIEELYREIGRVVYSKYENGEIIDKDLSDCCQQLGESYNNKI
```

```
IELEKDIDNIKSTKVNKKVKCDTDEDICTEYKSTEDSKIEDAKVSELHDSILKPED
DELVENCEGPGCNFTMTDGRDDDLIK"
complement(5249..5324)
/product="trNA-Lys"
/notes="codon recognized: AAA"
complement(5340..5427)
/product="5S ribosomal RNA"
complement(5606..8450)
/product="23S ribosomal RNA"
5938..6594
/locus_tag="CTC00065"
5938..6594
/locus_tag="CTC00065"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO34720.1"
/db_xref="GI:28202274"
/translacion="MRRADIEVNPLPVDVDSWGRSACYPRGSFYPLSDGPPTRNHRIT
KPDFPCSTCMCRSQAPLCCLYLIRASDRABGTGRLRYFLGGDRPSPQTAHLTWSRDQ
IHGPRLEFOYQGGIPRMTPOKLTLLLSLPPILYRQYRNSMLSYSKALRGLSLSRV
ASIFCTTISPDLLLRQCPSPHYAIRAGRNLPDKEFRYLRTVITVTAAGWGLSSPLRLR"
complement(8715..10223)
/product="16S ribosomal RNA"
10753..11397
/locus_tag="CTC00066"
10753..11397
/locus_tag="CTC00066"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AAO34721.1"
/db_xref="GI:28202275"
/translacion="MKNKNKILTIPTIIVIFAYNVKLVNDKNLSLPMNMDKSIEMQSR
KITRDVAVKKIQEKQDLITMEVBLNKNVILDSWGNLDIKFKYVISIDFYGTGIYSDR
LAKLNRENVLIQPSKTTIRIPKPKVKSITLNEKDTFTKTENGGFRFGREIKITPAESQ
ILSKAKERNQEQLNEKHLVESALSCTKSIEKSIQSIILLPYENYINVIKPIDG"
complement(12432..13208)
/locus_tag="CTC00067"
complement(12432..13208)
/locus_tag="CTC00067"
/codon_start=1
/product="putative polyferredoxin"
/protein_id="AAO34722.1"
/db_xref="GI:28202276"
/translacion="MRGVLCYFGSGTGNKMIADKFKYFNMKNIELDIINIIEEKQDME
LSEYNFLVVGTPYIGELPKIVDTFLNLPNKNKIKTILYSTQGTAKASALNVIESKI
ENKNYTVWSKIMFKMTNNFYISGKASKEIHEILISCERKVNKVTNNFNGEYKIKE
KVNKIRVTLSKIAKHPNNYLNHLSKNLSSENCEIRGCLCIRNCPQNNITFENGHVFI
HSCKMCLRCIYICPTNEIRYKNKIDQIQRIKSLKIK"
13557..14816
/locus_tag="CTC00068"
13557..14816
/locus_tag="CTC00068"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AAO34723.1"
/db_xref="GI:28202277"
/translacion="MKNLSIYNNYRIILLYFACFIIFTDPTISIFTLKNNIILQPD
ILITDYLEISGIGASFVNSATLCIYIILIRLAVKLSGSIVAGLFTVCGFSLFGKLN
LNVWPLFLGTWLYSKYQKEPPSKYILVALPGSTLSPAVSQISFSQHPETSGLIIGIS
LGIFMGFILPPISSYCLMHEGFNLNSGPAAGLVGTFPMSILRSFGINFORLWMST
KTYTLLIIFLPSILFLSMILVFCNLDSPFNKLDKLYKHSKPNKSDSYLIYIGEGSSFIN
MGILGIFMWYIILVIDGPLNGPIAGIFTIVGFGAGFKHFTIPVWLTGTIAGTINI
WSINSPSMLLATLFTSTLAPISGVFGPIYGIAGFLTCIVMNIHGLHAGLNLYNNGF
SGGFVALIILPIITNLKVDTONELQ"
complement(14871..15398)
/locus_tag="CTC00069"
complement(14871..15398)
/locus_tag="CTC00069"
```

```
/BC_number="3.6.1.22"
/codon_start=11
/transl_table=11
/product="predicted NADH pyrophosphatase"
/protein_id="AA034724.1"
/db_xref="GI:28202278"
/transl_table="MKVEYCPKGNKLIKESYSDGVPYCPIDDTMYRPTPKPCVVV
AVLKESQILLKQSYLPENSKVLYSVNGETVEETVREVKKEGIEVKNIKYLGS
EYLSKEIIMLIPFMAYSKSGVIDKSEVEVWNGHIEDALCQMSDEIGKNVVRKVLK
EIDYKGEKAYKCBIK"
complement (15446..16138)
/locus_tag="CTC00070"
complement (15446..16138)
/locus_tag="CTC00070"
/codon_start=1
/transl_table=11
/product="sugar fermentation stimulation protein"
/protein_id="AA034725.1"
/db_xref="GI:28202279"
/transl_table="MIKKNIKRABFIKPNRFOAVVVDNESIMVHPVNTGRCKEL
IPGTVLREESNPTKTYDLIGYKNGKFINIDSIQPNKVVEALKNGIKKJY
NIKREKTYGNRDFPLGNDGSEYLYEVKVTLEDGGVTFKFPDAPTERGTTHLEL
IDVKNKIGIGVLFQIMKQVKYFTPNIEMDKNFSEALVSAHKAGVDIPAYECVDVED
FITLKEVKVIL"
complement (16325..17551)
/locus_tag="CTC00071"
complement (16325..17551)

gene
CDS

Query Match      23.7%; Score 113.8; DB 1; Length 300029;
Best Local Similarity 56.3%; Pred. No. 3e-19;
Matches 245; Conservative 0; Mismatches 172; Indels 18; Gaps 1;

QY 32 AATTTTATAGAGCATATCTTACAAATCTACCTCACCGCTATCCCATCTTTTAGTGG 91
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 ATAGAATTATAGAGTTACAGCCATATAAAATTTGTCGCTTATAGAAATATCACTTTA 151
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92140 ATAAAGTAGAAAACAGTAGAACCTGGGAAAAAATTTGAGCATATAAAATGTAACATTA 92199
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 ATCAAGACGTTTAAACGGGCATTTCCCTAATAAGCCATTTTCCCGGGGTTTTCATCG 211
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92200 ATGAAGGATCTTTAGAGGACATTTCCCAACAGTAATGCGAGGAGTAAATTA 92259
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 TAGAGGCGATGGCGAAACAGGAGGCTTTTACCTTTCACTAGCTTGTGGGGGTTTGACC 271
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92260 TAGAGGCAATAGCTCAAGCGGGAGCTATAGCAGTATTAAGT-----T 92301
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 272 CTGAATCCGCAAAACAAAAATCGTGATTTTCATGACGATTTGATAAGGTTAAATTCGCA 331
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92302 TAGACGAATTTAAAGGTAATAATACCATCTTTGCTGGAATAAATAAGCTAAATTTAGAA 92361
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 TCCTGTATACCCGAGCGAGATTAAGATTAACATTTAGAAAGTCTTAAGCATAAGGGCA 391
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92362 AGAAAGTTATCCCAAGGAGTACTTTAAATTTAGAAAGTGAAGTAAAGTAAAGAGTT 92421
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 TGATCTCGCAAGTGGTGCACGGCTCAAGTGCATGCGCAAGTGGTCTGAAGCCGAAT 451
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92422 CAGCAGGTATAGCAAGGTATAGCTAAGGTGAATGAAAGTGGTAGCAGAAAGCAGAA 92481
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 452 TGAAGGCCATGATTG 466
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92482 TCATGTTTATGATAG 92496
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
AF124757
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
AF124757
Zymomonas mobilis fsmid clone 43D2, complete sequence.
AF124757.1 GI:4768846
Zymomonas mobilis
```

```
ORGANISM      Zymomonas mobilis
REFERENCE     Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
AUTHORS       Sphingomonadaceae; Zymomonas.
TITLE         Lee,H.J. and Kang,H.S.
JOURNAL       Sequence analysis of 43D2 fsmid clone of Zymomonas mobilis ZM4
REFERENCE     Unpublished
AUTHORS       2 (bases 1 to 34083)
TITLE         Lee,H.J. and Kang,H.S.
JOURNAL       Direct Submission
SUBMITTED     (29-JAN-1999) Microbiology, Seoul National University,
              San 56-1 Shillindong KwanakGu, Seoul 151-742, Korea
FEATURES     Location/Qualifiers
              1. 34083
                /organism="Zymomonas mobilis"
                /mol_type="genomic DNA"
                /strain="ZM4"
                /db_xref="taxon:542"
                /clone="fsmid 43D2"
                complement (106..11380)
                /gene="sah"
                complement (106..11380)
                /gene="sah"
                /codon_start=1
                /transl_table=11
                /product="succinylarginine dihydrolase"
                /protein_id="AAD29641.1"
                /db_xref="GI:4768847"
                /translation="MIVSENVFDLIGPTHNVAGLSRGNVASALHAGOPSYPQAAALQ
                GLEMKHMDMGLTQGVLPPLRPVTHLHLGVKDDKTLIKQAQDDRLALPNLCS
                ASSMWAANAATVISEPDSHGRVHPTANTLATMLRHLEAQTYYAQLNQMFNSCPFA
                MHPLPCQGHFSDGSAANHMRTISAHGRGTGNIIVYDKNDIYPARQKLRSQAVALR
                GEVYKDLAFIPQKEAIAKGFHNDVAVANVYLLAHAEAFEDQGWIKRIAEKID
                GFPIIDNTLEQAVKSYLFNSQIVTLKDRTMALILPQEVKSDPAVETVNRITISGN
                NPQIAVVDVRESMWANGGPACLRVPLSNALAEAVDQRFILTPKRWEMLYQLVEN
                FWDKTPDLDLVLDLTKTAHVAHWALTSWLG"
              1506..2540
                /note="hypothetical protein; ORF1"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="AAD29642.1"
                /db_xref="GI:4768848"
                /translation="MKPFVFIADTRLSDQEKLFQYEPAGYILFORNITDRTQLKK
                LTADLRANSQONTIPILIDQEGGRVWRMPPVPAAPFAARFDELYOKAPVSAIEAVRA
                NSHALALLDQVGINVDCMPLDIRDPEDGNIIGDRSGHNPEQVVALGRAVLGLAK
                GSVGCIKHIPGHQAKVDSKALPVDSPIELERDIRPQHLENAMKAMTAHLIYQ
                AWKESCASYSYKIIHDIIRKKGIGLMSDDIGMNALDGPVTRALRVFDAGCDVA
                LHCSGFDDMKALLENMDEIRPESFALEKVTNEFSASDFDLDRANKALPDRCIAE
                RDALFLEAAV"
                complement (2537..2845)
                /note="hypothetical protein; ORF2"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="AAD29643.1"
                /db_xref="GI:4768849"
                /translation="MNTAFDITITLAKADQEQALQALLHYADQSKTEGGLAFH
                IGDKSQPPFIILNRWQDHAAPFAHEKSDIFQDFASMFKFLKPLLSVFSVPLMED"
              3954..5909
                /gene="sah"
                complement (3954..5909)
                /gene="sah"
                /codon_start=1
                /transl_table=11
                /product="glutaryl-7-ACA acylase precursor"
                /protein_id="AAD29644.1"
                /db_xref="GI:4768850"
                /translation="MSRLKLSVNLKAMKYLVRGSSVVASLTSILAIPALSNAPDKTIW
                PENGDIAMHFSAPTADYKREVMIPMRQVKLHIVVFPNGNRLPILITRTPYDA
                SSRTSRSDSPMLATLPEGDEYVRDGYIRVFQDVRGKYGSEGVYVTRPVVDLNP
                KVDHTTDAWDTIEWLTKHPIESNVRGVMIGSSYEGFTVAALLNPHLPALKVAAAPSPM
```

```
VDGMGDDMFHYGAFROAFDYFLRQMTAKGTGAPPVHGAYDDYKAFLETSGAGNNAK
KEGDIQIPWQRULSIHPAYDFWQOQALDQILVAAPHSVPTWVQGLWDQEDMWGAIIH
SFEGLKVAHDTITNVLVMPGPNRWHSQVNYNGSLAGALHWDGDTALQFRDRDTLLPFNRY
LKDQKPAETKALITYNTGPNWKLNDWQTDKEKQLTPLYLAQASLSFTKPTSGNA
PSDQIISDPKPKVPYLPITPIFADTRWKMILIEDORFAASRPDVLTVETPLVDHPEK
LRGAPPANLLAATGSDVDWVVKLIDVYPDEIPSDPKMGQQLAISMDIFRGRIYNSF
EKSPSPAGKVCQQRFRFLPVVDHVFPLPGRHIMVQIQSSLFPLYDRNPQRYVENIFAK
PADYAAVTETVMHSPPDQASSVELFVP"
6171..7190
/gene="recA"
CDS
6171..7190
/gene="recA"
/function="recombinase"
/codon_start=1
/transl_table=11
/product="RecA protein"
/protein_id="AAD29645.1"
/db_xref="GI:4768851"
/translation="MDKQKALEAALQADRAFSGKSVMLRSREKIEIDTISGSLGL
DIALIGLGRIRIVEIYGPSSGKTTIALHAIABAKAGGTAAAFVDAEHALDPYAK
KLGVNTDOLIVSQPTGQALEITDILRINAVADILVDSVAALTPRAEIEGEMGDSH
VGLQARLMSQALRTKITSINRSOTLIVFINOVRMKIGVMYGNPETTGNALTKVASY
BLDIRVCQIKDRDEIVGNATRVKVKLLRPFKQVEFDIMYGGSVKMGELDLGV
KAGLIDKSGSWFSHDSVRIGORENAKTLREHPENTEKIEKMRHNTABIADEMDDP
SIPED"
7338..8114
/gene="comp"
CDS
7338..8114
/gene="comp"
/codon_start=1
/transl_table=11
/product="competence protein F"
/protein_id="AAD29646.1"
/db_xref="GI:4768852"
/translation="MPYRHAQIVARSFWLHSPFCVCSQSLHFLVPAACIOGPPLP
RDOQTEPCGLOHPSPFDRMNAAYVDEISRLILCKYKQKWLAKDMAKLMWRFL
LPISNIRKAVQNLLADHNSCPNNSSSPQTEPLIIPVLRWRLWQRFNGNQALITQY
LAKASGYDDCDLSLRHRSKTPGLHLTPKQHMVKNAPFLGIRKGRDLRGLDRIIFI
DDVTSGATTSKARLLKAGVSVHVLCAWVVIQDNFL"
complement(8200..8895)
/gene="phoB"
CDS
complement(8200..8895)
/gene="phoB"
/codon_start=1
/transl_table=11
/product="phosphate regulatory protein PhoB"
/protein_id="AAD29647.1"
/db_xref="GI:4768853"
/translation="MAALRLILVEDDDALAEILLKWHFSKQNFVHVHTRDGEALLLAR
ENPPDMLDMWVEGTSGIEVCQRLRRYPETANLPIIMLTARGEEDDRIRGLETGADD
YVTKPSPRELVAFLVLRIRPALSKEQALNYNDLMDLVSHRVKRACTNVSIGPTE
FALLRHMEYFRVRSREKLLDIYIWQDSIDIELRTVDVHIRLRKALNRDGLPDLIRT
VRSAGYALDTDEE"
complement(8895..9608)
/gene="phoU"
CDS
complement(8895..9608)
/gene="phoU"
/codon_start=1
/transl_table=11
/product="transcriptional regulator PhoU"
/protein_id="AAD29648.1"
/db_xref="GI:4768854"
/translation="MKKASEIGHYKAFDEDLGLRALICEMGGLTEIALTSTLEALI
HSNKELAAQIVSEKEDIALEAEARVAMRIISLRAPLADDLRDVTAANKISGILERM
ADYAKNIARVPILQNMVSIHPVPLVPMGEMVSMVRNVLDAYAADPKKALHVAEY
DMEVDGLYNLFRILLTYMMEDSGRISTCIHLMFIAKRLERIGDHATNIAEMVYFAAT
GKQMPERTRGDILLASEGS"
complement(9815..11080)
/gene="phoR"
CDS
complement(9815..11080)
/gene="phoR"
/codon_start=1
/transl_table=11
```

```
/product="phosphate regulon sensor protein PhoR"
/protein_id="AAD29649.1"
/db_xref="GI:4768855"
/translation="MTFPQSRLEKCKVICRIPVYLSGLSIALFIIVSHSVSGMPIL
SLMLIVILGMWPKNISEASTPPKESGNSLTDPAVPSQATAADIIIAMDEPALLV
NCGOVAVNAAENLLGMHIEGGDIRMAIRHPAAIRLLTRELDMNTVLLAGLGQVNR
RVELIYAPINEQRLLILLRDOQSTAHLTQEVRIIDPVANTSEHLRTPLATLSGFIELED
VDKIKERDRHHFELSIMAREAKQMNLDLMSLSRTEAGKFSPLHDIIVHMSPVVEEV
VNIQASQAKATQITLENHLSDSLQSDRAELTQLLYNTIGNALKYRKDGI1KVAL
DSNTDHRIKLAIADEGDGIPIYHHPRLTERFYRVDSRSRRLGCTGLGLAIYKQIVEH
HRGELLIDSLPGKGTVTVPFLTTFDPE"
11284..12006
/note="ORF3"
CDS
11284..12006
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAD29650.1"
/db_xref="GI:4768856"
/translation="MNKEQRKNSAQSIRWHPHSIHPEGRKFFVIAVIIISALITYFS
WLMWAMPLGVLCFAAFPRDPIRTVPEGGLIVSPADGMVCLLADVPPPELAGADG
LGDAPLTRVSIIFMSVDFHINRAPVRIARIAYIPGAFVNADLDKASEKNERQHFI
ETKEGINVGLTQIAGLIARRIVPLTKVEDYVERGERICLIRFGSRLDYLPAGVTPQV
ALGQRCLAGTETWDLRDKPSR"
12147..12821
/gene="psaA"
CDS
12147..12821
/gene="psaA"
/codon_start=1
/transl_table=11
/product="CDP-diacylglycerol--serine
O-phosphatidyltransferase"
/protein_id="AAD29651.1"
/db_xref="GI:4768857"
/translation="MRYPVWEHAIVAILVAALVDLDRVARLLKGATRGAEILDSL
SDAIAQVAPAILVYLSLQNPWRGWLFALSHAVCCALRLARFNVNIDYLDOPKSA
GFUTGPAPAGNFSWMPPLVANTVSYVDLFRMPELLVGLWLSLAFMLADLPTGMSV
LRVRPKLRUGLLVYVAGAAAFISAPWAVLSVFSVIYTMDFPGLISYSKIMKSKEKT
IVAEES"
complement(12976..13674)
/gene="asp"
CDS
complement(12976..13674)
/gene="asp"
/codon_start=1
/transl_table=11
/product="aspartate racemase"
/protein_id="AAD29652.1"
/db_xref="GI:4768858"
/translation="MKKGLIGIGPESTLLYYKLVYGANNRIGSQYFPLTVBSLN
Query Match 22.4% Score 107.6; DB 1; Length 34083;
Best Local Similarity 55.7% Pred. No. 1.4e-17;
Matches 234; Conservative 0; Mismatches 174; Indels 12; Gaps 1;
Qy 60 TCTACCTCAGCGGTATCCCATGCTTTTAGTGATAGATAATATTAGAGTTACAAGCCATAA 119
Db 25278 TCTGCGCTCACGGTTTTCGGATGCTGCTGGTGTGATCGGGTCGAGACACTGATTCGGGCGA 25337
Qy 120 AAAAAATTGTCCTTATAAGAAATACACTTTTAAAGAGACGCTGTTTAAACGGGCATTTC 179
Db 25338 AAAAAATTGTCGATCAAGGCCGTAACAAATTAACGAGCGGTCCTTTACCGTCATTTC 25397
Qy 180 TAATAAGCCCATTTTCCCGGCGCTTTTGATCGTAGAGGCATGGCGCAACCGGAGGGTT 239
Db 25398 CGGTAAATCCGATTATCCAGGTGTTTTAAATGTGCGAGCCCTGGCACAAGCTGCCGTT 25457
Qy 240 TTTAGCCTTCACAGCTTGTGGGGTGTGACCCCTGAATCGCCAAAAACAAAATCGTGTGA 299
Db 25458 TCTTGCCATTGAAAGCTTAGGACTG-----ACCGGAACCGGTAAAGCTGTTTA 25505
Qy 300 TTTTCATGACGATTTCATAAGGTTAAATTTCCGCATCCCTGTAAACCCAGGCGACAGATTAGA 359
Db 25506 TTTTCATGCGCTATTAAACGAGACAAAATTCGCTATTCCCGGTAGAACCGGGTATTTTATTGCG 25565
```

QY 360 ATACCATTTAGAACTCTTAAGCATTAAGGCGCATGATCTGCGAAGTGGTGGCGACGGCTCA 419  
Db 25566 GTTAGAAGTCAGGTTTTTACAAAACGGGCAAGAATTTGAATTTAAAGCCCGTGCCCTT 25625  
QY 420 AGTGGATGCGCAAGTGGTCTGAAAGCCGAATTTGAAGCCATGATTGCGAGAGAGATTA 479  
Db 25626 GATTGAAGACAAAGTCGGGCTGAACAGAGATTTACCGCCATGATCGCTGATCCAGCTAA 25685

## RESULT 12

AE008692.11/c

## WPCOMMENT

Sequence split into 21 fragments LOCUS AE008692 Accession AE008692

Fragment Name Begin End

AE008692\_00 1 110000

AE008692\_01 100001 210000

AE008692\_02 200001 310000

AE008692\_03 300001 410000

AE008692\_04 400001 510000

AE008692\_05 500001 610000

AE008692\_06 600001 710000

AE008692\_07 700001 810000

AE008692\_08 800001 910000

AE008692\_09 900001 1010000

AE008692\_10 1000001 1110000

AE008692\_11 1100001 1210000

AE008692\_12 1200001 1310000

AE008692\_13 1300001 1410000

AE008692\_14 1400001 1510000

AE008692\_15 1500001 1610000

AE008692\_16 1600001 1710000

AE008692\_17 1700001 1810000

AE008692\_18 1800001 1910000

AE008692\_19 1900001 2010000

AE008692\_20 2000001 2056415

Continuation (12 of 21) of AE008692 from base 1100001 (AE008692 Zymomonas mobilis subsp.

## Query Match

Best Local Similarity 22.4%; Score 107.6; DB 1; Length 110000;

Matches 234; Conservative 55.7%; Pred. No. 1.4e-17;

Matches 234; Mismatches 0; Indels 12; Gaps 1;

QY 60 TCTACCTACCGCTATCCGATGCTTTTATGAGTAAATATAGATTACAGCCATATA 119

Db 63029 TCTGCTCCTACCGCTTTTCCGATGCTGCTGATCGGTCGAGACACTGATTCGGGCGCA 62970

QY 120 AAAAATTGTCGCTTATAAGAAATACACTTTTAAAGAGCGTGTAAACGGCATTTCCC 179

Db 62969 AAAAATTGTCGATCAAGCCGCTAACAAATTACGAGCGCTTTTACCGGCTCATTTCCC 62910

QY 180 TAATAAGCCCATTTTCCCGGGGCTTTTATGATCGTAGAGGCGATCGCGCAACGGGAGGGTT 239

Db 62909 CGGTAATCCGATATGCGAGGTGTTTAAATGTCGAGGCTTGGCACAGCTGCCGGTGT 62850

QY 240 TTTAGCCTTCACTAGCTTGTGGGGGTTTGACCCCTGAAATGCGCAAAACAAAATCGTGA 299

Db 62849 TCTTGCAATTGAAAGCTTAGGACTG-----ACCGGAACCGGTAAAGCTGGTTTA 62802

QY 300 TTTTCATGAGATTGATAAGTTAAATTCGCTATCCCTGTTAAACCCAGGCGGACAGATTAGA 359

Db 62801 TTTTCATGAGTTAATACGAGACAAAATTCGCTATTCCGGTAGAACCGGGTATTTTATTGG 62742

QY 360 ATACCATTTAGAACTCTTAAGCATTAAGGCGCATGATCTGCGAAGTGGTGGCGACGGCTCA 419

Db 62741 GTTAGAAGTCAGGTTTTTACAAAACGGGCAAGAATTTGAATTTAAAGCCCGTGCCCTT 62682

QY 420 AGTGGATGCGAAGTGGTCTGAAAGCCGAATTTGAAGCCATGATTGCGAGAGAGATTA 479

Db 62681 GATTGAAGACAAAGTCGGGCTGAACAGAGTTCCACCGCCATGATCGCTGATCCAGCTAA 62622

## RESULT 13

AE017180.24/c

## WPCOMMENT

Sequence split into 39 fragments LOCUS AE017180 Accession AE017180

Fragment Name Begin End

AE017180\_00 1 110000

AE017180\_01 100001 210000

AE017180\_02 200001 310000

AE017180\_03 300001 410000

AE017180\_04 400001 510000

AE017180\_05 500001 610000

AE017180\_06 600001 710000

AE017180\_07 700001 810000

AE017180\_08 800001 910000

AE017180\_09 900001 1010000

AE017180\_10 1000001 1110000

AE017180\_11 1100001 1210000

AE017180\_12 1200001 1310000

AE017180\_13 1300001 1410000

AE017180\_14 1400001 1510000

AE017180\_15 1500001 1610000

AE017180\_16 1600001 1710000

AE017180\_17 1700001 1810000

AE017180\_18 1800001 1910000

AE017180\_19 1900001 2010000

AE017180\_20 2000001 2110000

AE017180\_21 2100001 2210000

AE017180\_22 2200001 2310000

AE017180\_23 2300001 2410000

AE017180\_24 2400001 2510000

AE017180\_25 2500001 2610000

AE017180\_26 2600001 2710000

AE017180\_27 2700001 2810000

AE017180\_28 2800001 2910000

AE017180\_29 2900001 3010000

AE017180\_30 3000001 3110000

AE017180\_31 3100001 3210000

AE017180\_32 3200001 3310000

AE017180\_33 3300001 3410000

AE017180\_34 3400001 3510000

AE017180\_35 3500001 3610000

AE017180\_36 3600001 3710000

AE017180\_37 3700001 3810000

AE017180\_38 3800001 3814139

Continuation (25 of 39) of AE017180 from base 2400001 (AE017180 Geobacter sulfurreducens

Query Match 22.2%; Score 106.4; DB 1; Length 110000;

Best Local Similarity 53.9%; Pred. No. 2.8e-17;

Matches 249; Conservative 0; Mismatches 201; Indels 12; Gaps 1;

QY 19 AACTTGAATCTCAATTTTTTATAGAGCATATCTTACAAATTTCTACCTCACCGCTATCCC 78

Db 81306 AACATGGAACAGTCTTTTGATATCAACGAATCATGAAGATTTCTCCCGCACCGCTACCCC 81247

QY 79 ATGCTTTTAGGTAGATAGATTATAGATTACAGCCATTAATAAATTTGTCGTTATAG 138

Db 81246 TTCTCTCTGTGGATCGGATTTGGAGTACGTTCCGCGGACGGGATCGTGGGGATCAAG 81187

QY 139 AATATCACTTTTAATGAAGACGCTGTTAAACGGGCATTTCCCTAATAAGCCATTTTCCCG 198

Db 81186 AAGTCAGCATCAATGAGCCGTTTTTCCAGGCCACTTTCCGGGACACCCGGTATGCG 81127

QY 199 GCGTTTTTGTATCGTAGAGGCGATCGCGCAACACGGAGGGTTTTTATAGCCTTCACTAGCTTG 258

Db 81126 GGAGTCTCTGATCGTTGAGGCCATGCGCCAGTGGGGGAATCTAGCCCTATGTGACCTTG 81067

QY 259 TGGGGGTTTGAACCTGAATTCGCCCAAAACAAAATCGTGTATTTTCATGACCATGATAG 318

Db 81066 GCGGACGAGGTCCGCGA-----CAAGGCTCTGCTACTTTTGCCTCCATCGACAAC 81019

QY 319 GTTAAATTCGCGCATCCCTGTATAACCCAGGCGACAGATTAGAAATACCAATTTTGAAGTCTTA 378

Db 81018 GTCAAGTTCCGCAAAACCGGTGGTCCCGGAGACCAACTGCGGATCGAAGTACGATCTCC 80959

QY 379 AAGCATAAAGGCGCATGATCTTGGCAAGTGGGTGGCACGGCTCAAGTGGATGGCAAGTGGTC 438



Db 80958 GGTTCACAGCGGGGCACTGTGGTCTTACGCCGAGGGCCACCGTGAACGGCAAGCTGGTC 80899  
 QY 439 GCTGAAGCCGAATTGAAGCCCAATGATTGCGAGAGAGATTAA 480  
 Db 80898 ACCGAGCGGAGCTCAAGCGGACCTTTGCGGCAAGGAAAAA 80857  
  
 RESULT 14  
 BX470161  
 LOCUS  
 DEFINITION Zebrafish DNA sequence from clone RP71-8412, complete sequence.  
 ACCESSION BX470161  
 VERSION HTG.  
 KEYWORDS Danio rerio (zebrafish)  
 SOURCE Danio rerio  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 197615)  
 Hunter, G.  
 Direct Submission  
 Submitted (31-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Jul 31, 2004 this sequence version replaced gi:35209644.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Web site: http://www.sanger.ac.uk  
 Contact: zfish-help@sanger.ac.uk  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest, except on the rare  
 occasion of the clone being a YAC.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
 Zebrafish pUC subclones occasionally display inconsistency over the  
 length of mononucleotide A/T runs and conserved TA repeats. Where  
 this is found the longest good quality representation will be  
 submitted.  
 Repeat names beginning 'Dr' were identified by the Recon repeat  
 discovery system (Zhiron Bao and Sean Eddy, submitted), and those  
 beginning 'dir' were identified by Rick Waterman (Stephen Johnson  
 lab, WashU). For further information see  
 http://www.sanger.ac.uk/Projects/D\_rerio/fishmask.shtml RP71-8412  
 is from a Zebrafish mixed sex BAC library VECTOR: pTARBAC2.  
  
 FEATURES  
 source  
 Location/Qualifiers  
 1..197615  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="RP71-8412"  
 /clone\_lib="RP71-8412"  
  
 Query Match 21.6%; Score 103.8; DB 5; Length 197615;  
 Best Local Similarity 54.6%; Pred. No. 1.3e-16;  
 Matches 241; Conservative 0; Mismatches 182; Indels 18; Gaps 1;

QY 40 ATAGAGCATATCTTACAAATTTACCTACCGCTATCCCATGCTTTTAGTGATAGATT 99  
 Db 162780 ATCTTGAGATCATGGAACCTGCTGCGCATCGTTATCCGTTCTGCTGTCGCAAGATC 162839  
 QY 100 ATAGAGTTACAAGCCCAATAAAAAATTTGTCGCTTATAAGAAATATCACATTTTAAATGAAGAC 159  
 Db 162840 ATCGAGATCAAAAGTGACGATCGGCCCATCGGATCAAGAAATGTGACGGCCAATGAGCGG 162899  
 QY 160 GTGTTTAAACGGGCATTTCCCTTAATAAGCCCATTTTCCCGGGCGTTTTCATCGTAGAGGCG 219  
 Db 162900 CAATTCTCTGGGATTTCCCTTGGGACCGCATCATGCGGGCGTGTCTCTGTCGAGGCG 162959  
 QY 220 ATGGCGCAACCGGAGGGTTTTTAGCCCTTCACTAGCTTGTGGGGGTTTGACCCCTGAATATC 279  
 Db 162960 ATGGCCGACAGCGCGCGGCATCTGCGCCGCAAGCAGGG-----C 163001  
 QY 280 GCCAAACAAAATTCGTATTTTCATGACGATTTGATAGGTTAAATTTCCGCATCCCTCTGA 339  
 Db 163002 CGCAGCGGCAATCTGGTCTACTTCATGACCATGACAAACGCCCGGTTCGCAAGCCGTC 163061  
 QY 340 ACCCCAGGCGCAGATTAGAATACCAATTAGAAGTCTTAAAGCATAGGGGCATGATCTGG 399  
 Db 163062 GTCCGGCGCACCGCTGGAAATATCATGCTCAAGCAGAGCAGCGTGGCAATATCTGG 163121  
 QY 400 CAAGTGGGTGGCACGGCTCAAGTGGATGGCAAGTGTGCTGAAGCGCAATTTGAAAGCC 459  
 Db 163122 AAGTTCATTTGCGACGCCAAGGTGGACGGCGCATTTGTTGGCGAAGCCGATTTTCGGTGG 163181  
 QY 460 ATGATTCACAGAGAGATTAA 480  
 Db 163182 ATGATGTGAACAAGGACGAA 163202  
  
 RESULT 15  
 SMES91787  
 LOCUS  
 DEFINITION Sinorhizobium meliloti 1021 complete chromosome; segment 6/12.  
 ACCESSION AL591787 AL591688  
 VERSION AL591787.1 GI:15074266  
 KEYWORDS  
 SOURCE Sinorhizobium meliloti (Rhizobium meliloti)  
 ORGANISM  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 1 (bases 1 to 329100)  
 Capela, D., Barloy-Hubier, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J.,  
 Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreano, S., Gloux, S.,  
 Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D.,  
 Pohl, T., Portetle, D., Puhler, A., Purnelle, B., Ramsperger, U.,  
 Renard, C., Thebault, P., Vandenbol, M., Weidner, S. and Galibert, F.  
 Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021  
 JOURNAL  
 Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001)  
 PUBMED  
 11481430  
 REFERENCE  
 2 (bases 1 to 329100)  
 Gouzy, J.  
 Direct Submission  
 Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO  
 EU Consortium  
 COMMENT  
 MELILO EU Consortium:  
 Laboratoire de Biologie Molculaire des Relations  
 Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet,  
 France, Laboratoire de Genetique et Developpement UMR6061-CNRS,  
 Faculte de Medecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes,  
 France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany,  
 Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25,  
 D-33615 Bielefeld, Germany, Unite de Biochimie physiologique,  
 Universite Catholique de Louvain, Place Croix du Sud 2, Bte 20,  
 B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculte  
 des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6,  
 B-5030 Gembloux, Belgium. E-mail:Jerome.Gouzy@toulouse.inra.fr  
 http://sequence.toulouse.inra.fr/meliloti.html.

## FEATURES

source	Location/Qualifiers
gene	1. .329100 /organism="Sinorhizobium meliloti" /mol_type="Genomic DNA" /strain="1021" /db_xref="taxon:382" 152. .227 /gene="tRNA-THR_TGT"
tRNA	132. .227 /gene="tRNA-THR_TGT" /product="tRNA-Thr" /note="codon recognized: ACA; predicted by tRNAscan-SE" /evidence=not experimental complement(373. .798) /gene="SMc04435" complement(373. .798) /gene="SMc04435" /function="miscellaneous; hypothetical/global homology" /note="Product confidence : hypothetical Gene name confidence : hypothetical predicted by Homology" /codon_start=1 /evidence=not experimental /transl_table=11 /product="CONSERVED HYPOTHETICAL PROTEIN" /protein_id="CAC45913.1" /db_xref="GI:15074267" /db_xref="InterPro:IPR004360" /db_xref="IPR011588" /db_xref="UniProt/TREMBL:Q92Q19" /translation="MNNALVPFVAVSDWQAKSEFYCNILGFCSDYERPEEGCYLSLD GAELMDIGTGRTPAGLHLPSPYFPGRLNVIQIRVASADALVRALTDRGIALYPLP ERWYRGDEGVGNVPADPDGYLLRFYELGRPRPTS" complement(914. .1855) /gene="SMc01330" complement(914. .1855) /gene="SMc01330" /function="miscellaneous; not classified regulator" /note="Product confidence : putative Gene name confidence : hypothetical predicted by Codon usage predicted by Homology predicted by Framed" /codon_start=1 /evidence=not experimental /transl_table=11 /product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN" /protein_id="CAC45914.1" /db_xref="GI:15074268" /db_xref="GOA:Q92Q18" /db_xref="InterPro:IPR000847" /db_xref="InterPro:IPR005119" /db_xref="InterPro:IPR011991" /db_xref="UniProt/TREMBL:Q92Q18" /translation="MQDLNLTALFAAVVYKHGFSAAARDLNVPKSLKHVARLBEQL GVRLERTRKLWTEGRI FYEHAQGLDGVAAAEIAA VRAEPTGVVRLACPIGF TPMLADILPAFRYPGVRLAITASNRIDILIEERIDVALRARDOLDTDSOLIVKFG EVRGLAASPTLLARGLHTDNLDSQMTLSNNEQHPNDVRLVLTGGAIAIAHRPV VGCSDFLLERAALEGMIALLPDHICERAFRTGALVPVLPWTSGNVVWHLVFPFSRH GLLPATRALIDFLAENLIKALERCREVDPRAASPEI" 2063. .2683 /gene="acpD" /note="synonym: SMc01329" 2063. .2683 /gene="acpD" /function="small molecule metabolism; fatty acid biosynthesis" /note="Product confidence : probable Gene name confidence : putative predicted by Codon usage predicted by Homology predicted by Framed" /codon_start=1
gene	/evidence=not experimental /transl_table=11 /product="PROBABLE ACYL CARRIER PROTEIN PHOSPHODIESTERASE" /protein_id="CAC45915.1" /db_xref="GI:15074269" /db_xref="GOA:Q92Q17" /db_xref="InterPro:IPR003680" /db_xref="UniProt/Swiss-Prot:Q92Q17" /translation="MKILHDSGILGSHSVRRRLTSAIVSQLKADRPDAEITYRDLAS LRVPHLTGAQIMAPADLEGVDALLAADVRIGRQMLEFLAADTVVVGAPMNFSLPSQ LKAWIDRLAVAGKTFRYTEAGPEGLAKGKLLIVASTRGHYSVAPASAMDHQETYLRS VFQFFGTDIEFTRAEGNLGPDQKQPAIAEAETIAEGNVKLAS" 2792. .3076 /gene="SMc01328" 2792. .3076 /gene="SMc01328" /function="miscellaneous; hypothetical/partial homology" /note="Product confidence : hypothetical Gene name confidence : hypothetical predicted by Codon usage predicted by Framed" /codon_start=1 /evidence=not experimental /transl_table=11 /product="HYPOTHETICAL TRANSMEMBRANE PROTEIN" /protein_id="CAC45916.1" /db_xref="GI:15074270" /db_xref="GOA:Q92KE9" /db_xref="UniProt/TREMBL:Q92KE9" /translation="MKCVRLRPLALSALAVALVAGSVASAPGRCCTCRHDDGAKYE LQGTACIRIGDISYLARCENNLVTKRILQDGCPTAGLAVTVIFSAPIP" complement(3066. .3941) /gene="SMC01327" complement(3066. .3941) /gene="SMC01327" /EC_number="2.1.1.1" /function="macromolecule metabolism; macromolecule synthesis, modification; rna synthesis, modification , dna transcription" /note="Product confidence : putative Gene name confidence : hypothetical predicted by Codon usage predicted by Homology predicted by Framed" /codon_start=1 /evidence=not experimental /transl_table=11 /product="PUTATIVE TRNA/RRNA METHYLTRANSFERASE PROTEIN" /protein_id="CAC45917.1" /db_xref="GI:15074271" /db_xref="GOA:Q92Q16" /db_xref="InterPro:IPR001537" /db_xref="InterPro:IPR004441" /db_xref="UniProt/TREMBL:Q92Q16" /translation="MSKDKTGKSARDTHYANLRRAHRDARRERGEIPTPRQDKRAKL SADWKPALAPQVLLYGLHVTAAALDNPERRIIRLSVTQNAARLDGLDTALPFPPT ETVMPQLDRILGPDALHOGWMLESTOPLPHRKEALRDCPLILVLDVOTDPHNVGAIM RAVAFDAPGLITTRHSPSTESGLAKSAQLELPIQITNLADTLBELHGLFMT IGLDSGPOPLEGTFFAGDKIALVLGSEKGLRQKRTQTVKALARLDMPGAIKSLNVS AAAAIYAAARHLKG" 4158. .4242 /gene="tRNA-TYR_GTA" 4158. .4242 /gene="tRNA-TYR_GTA" /product="tRNA-Tyr" /note="codon recognized: UAC; predicted by tRNAscan-SE" /evidence=not experimental 4266. .4339 /gene="tRNA-GLY_TCC" 4266. .4339 /gene="tRNA-GLY_TCC" /product="tRNA-Gly" /note="codon recognized: GGA; predicted by tRNAscan-SE"
gene	
tRNA	
gene	
tRNA	



```
gene /evidence=not_experimental
4602..5777
/gene="tufB"
/note="synonym: SMC01326"
CDS 4602..5777
/gene="tufB"
/function="macromolecule metabolism; macromolecule
synthesis, modification; proteins and peptides -
translation and modification"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon usage
predicted by Homology
predicted by FrameD"
/codon_start=1
/evidence=not_experimental
/transl_table=11

Query Match 21.6%; Score 103.8; DB 1; Length 329100;
Best Local Similarity 56.7%; Pred. No. 1.3e-16;
Matches 242; Conservative 0; Mismatches 167; Indels 18; Gaps 2;

QY 40 ATAGAGCATATCTTACAAATTTCTACCTCACCCTATCCCATGCTTTTGTGGATAGAAAT 99
Db 165453 ATACAGGAATCCTGAGGCTTCTTCCCATCGTTACCCCTTCTGCTGGTGGACCGCATC 165512

QY 100 ATAGAGTTACAAGCCAAATAAAAAAATTTGTCGCTTATAAGAAATATCACTTTTAATGAAGAC 159
Db 165513 ATCGAGATCGATGACGACAAATTCGGCCATCGGGATCAAGAACGTGACGGCCACGAGCCG 165572

QY 160 GTGTTTAAAGGGCATTTCCCTAATAAGCCCAATTTCCCGGGGCTTTTGATCGTAGAGGGC 219
Db 165573 CATTTTCAACGGGCACATTTCCCGAGAGCCGATCATCGCGGCGTTCTTCTGATCGAAGGC 165632

QY 220 ATGGCGCAACCGGAGGGTTTTTAGCCTTCACTAGCTTTGTTGGGGTTTGACCTGAAATC 279
Db 165633 ATGGCGCAGACGGCGCG-----CGCGATCTGTGCGGCAAGACGGGAATC 165677

QY 280 GCCAAACAAAATCGTGATTTTCATGACGATTGATAAGTTAAATTCGCGATCCCTGTA 339
Db 165678 G---GAAGCAATCTCGTCTATTTTCATGACGATCGCAATGCACGTTCCGCAACCCGTC 165734

QY 340 ACCCCAGGCGACAGATTAGAAATACCAATTTAGAACTTTAAAGCATAAAGGGCATGATCTGG 399
Db 165735 GTGCCGGGCGACCGGTGGAGTTCCATGTCGTGAAGCAGAGCAGCGCGGCAATATCTGG 165794

QY 400 CAAAGTGGTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTCTGCTGAAGCCGAATTTGAAGCC 459
Db 165795 AAATTTCACTGGCATGCAAAAGTTGACGGGCAACTGTCGCGGAAGCTGATATCGGCGCG 165854

QY 460 ATGATTG 466
Db 165855 ATGATTG 165861
```

Search completed: January 12, 2006, 02:20:52  
Job time : 3018 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2006, 21:53:19 ; Search time 461 Seconds  
(without alignments)  
6939.376 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggacaaagccatcaaaaa.....tgattgcagagagatttaa 480

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 496997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	100.0	480	8	ACA34928
2	454.4	94.7	480	4	AAS53867
3	294	61.3	853	6	ABX66131
4	134.4	28.0	441	8	ACA30215
5	100	20.8	659	6	ABX66187
6	99.2	20.7	459	13	ADS59846
7	95	19.8	432	8	ACA28687
8	93.4	19.5	417	8	ACA33871
9	92.6	19.3	462	8	ACA53362
10	88.6	18.5	435	8	ACA36484
11	88.6	18.5	1755	6	ABQ69109
12	88.6	18.5	1927	6	ABQ70710
13	88.6	18.5	110000	6	ABA03041_25
14	88.6	18.5	110000	6	ABA03041_26
15	88.2	18.4	459	8	ACA43412
16	87.4	18.2	453	8	ACA26987
17	87	18.1	110000	6	ABQ67196_5
18	87	18.1	110000	6	ABQ69245_26
19	86.6	18.0	453	10	ACF69720

C	20	86.6	18.0	110000	10	ACF67367_27	Continuation (28 o
	21	86.6	18.0	110000	10	ACF65386_1	Continuation (2 of
	22	86.4	18.0	426	4	AAS53189	Aas53189 Enterococ
	23	86.4	18.0	429	10	ADH83263	Adh83263 Enterococ
	24	86.4	18.0	6021	2	AAX13395	Aax13395 Enterococ
	25	86.4	18.0	6021	6	ABS99190	Abs99190 Enterococ
	26	86.4	18.0	243335	14	ABE42735	AbE42735 L. pneumo
	27	86.4	18.0	298667	14	ABE39173	ABE39173 L. pneumo
	28	86.2	18.0	414	13	ADS63118	AdS63118 Bacterial
	29	86.2	18.0	414	13	ADS62218	AdS62218 Bacterial
	30	86.2	18.0	414	13	ADS62656	AdS62656 Bacterial
	31	86.2	18.0	414	13	ADS59716	AdS59716 Bacterial
	32	86.2	18.0	426	8	ACA27788	ACA27788 Prokaryot
	33	85	17.7	423	8	ACA33329	ACA33329 Prokaryot
	34	85	17.7	471	13	ADS45464	AdS45464 Bacterial
	35	84.8	17.7	518	13	ADT46454	AdT46454 Bacterial
	36	84.8	17.7	110000	14	AEBS39175_06	Continuation (7 of
	37	84.8	17.7	295644	14	AEBS35721	AEb35721 L. pneumo
	38	83.6	17.4	420	13	ADT45905	AdT45905 Bacterial
	39	83.2	17.3	110000	14	AEBA42401_05	Continuation (6 of
	40	82.8	17.2	438	8	ACA45733	ACA45733 Prokaryot
	41	82.8	17.2	441	4	AAS4202	Aas4202 Pseudomon
	42	82.8	17.2	441	8	ACA42462	ACA42462 Prokaryot
	43	82.8	17.2	441	10	ADE49810	Ade49810 (3R)-hydr
	44	82.8	17.2	441	10	ADE49808	Ade49808 (3R)-hydr
C	45	82.8	17.2	465	11	ABD05628	ABd05628 Pseudomon

## ALIGNMENTS

RESULT 1  
ACA34928

ID ACA34928 standard; DNA; 480 BP.

XX ACA34928;

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #16585.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;  
drug design; gene.

OS Helicobacter pylori.

XX WO20027183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.  
P-PSDB; ABU31058.

XX New antisense nucleic acids, useful for identifying proteins or screening  
for homologous nucleic acids required for cellular proliferation to  
isolate candidate molecules for rational drug discovery programs.

PT Claim 14; SEQ ID NO 22798; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 480 BP; 150 A; 93 C; 107 G; 130 T; 0 U; 0 Other;

Query Match 100.0%; Score 480; DB 8; Length 480;  
Best Local Similarity 100.0%; Pred. No. 1.1e-137;  
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGAAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
1 ATGGAAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
61 CTACCTCACCCTATCCATGCTTTTATAGAGTATAGATTAAGGCAATATAA 120  
61 CTACCTCACCCTATCCATGCTTTTATAGAGTATAGATTAAGGCAATATAA 120  
121 AAAATTGCTGCTTATAGATATACATTTTATAGAGAGCTTTTAAACGGGCAATTCCT 180  
121 AAAATTGCTGCTTATAGATATACATTTTATAGAGAGCTTTTAAACGGGCAATTCCT 180  
181 AATAAGCCATTTTCCGGCGTTTGTATGATAGAGGCGATGCGCAACGGGAGGTTT 240  
181 AATAAGCCATTTTCCGGCGTTTGTATGATAGAGGCGATGCGCAACGGGAGGTTT 240  
241 TTAGCCTTCACTAGCTTTGGGGGTTTGACCTGAAATGCCAAACAAATAATCGTGAT 300  
241 TTAGCCTTCACTAGCTTTGGGGGTTTGACCTGAAATGCCAAACAAATAATCGTGAT 300  
301 TTATGACGATTAAGTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 360  
301 TTATGACGATTAAGTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 360  
361 TACCATTTAGAGCTTTTAAAGTATAGGCGATCATCTGGCAAGTGGTGGCGGCTCAA 420  
361 TACCATTTAGAGCTTTTAAAGTATAGGCGATCATCTGGCAAGTGGTGGCGGCTCAA 420  
421 GTGGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480  
421 GTGGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480

RESULT 2  
AAS53867  
IP AAS53867 standard; DNA; 480 BP.

XX AAS53867;  
XX 13-FEB-2002 (first entry)  
XX Helicobacter pylori DNA for cellular proliferation protein #321.  
XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;  
XX antibacterial; drug design.  
XX Helicobacter pylori.  
XX WO200170955-A2.  
XX 27-SEP-2001.  
XX 21-MAR-2001; 2001WO-US009180.  
XX 21-MAR-2000; 2000US-0191078P.  
XX 23-MAY-2000; 2000US-0206848P.  
XX 26-MAY-2000; 2000US-0207727P.  
XX 23-OCT-2000; 2000US-0242578P.  
XX 27-NOV-2000; 2000US-0253625P.  
XX 22-DEC-2000; 2000US-0257931P.  
XX 16-FEB-2001; 2001US-0269308P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
XX Yamamoto RT, Xu HH;  
XX WPI; 2001-611495/70.  
XX P-PSDB; AAU36008.  
XX New polynucleotides for the identification and development of  
XX antibiotics, comprise sequences of antisense nucleic acids.  
XX Claim 27; SEQ ID NO 7504; 511pp; English.  
XX The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the genes,  
XX their use in the discovery of novel antibiotics, the essential genes,  
XX themselves and the encoded proteins. The prokaryotes used are *Escherichia*  
XX *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
XX *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
XX useful for the identification of potential new targets for antibiotic  
XX development. The antisense nucleic acids can also be used to identify  
XX proteins used in proliferation, to express these proteins, and to obtain  
XX antibodies capable of binding to the expressed proteins. The proteins can  
XX be used to screen compounds in rational drug discovery programmes. The  
XX antisense nucleic acid sequence is also useful to screen for homologous  
XX nucleic acids which are required for cell proliferation in a wide variety  
XX of organisms. The present sequence encodes an essential prokaryotic  
XX cellular proliferation protein. Note: The sequence data for this patent  
XX did not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 480 BP; 151 A; 94 C; 107 G; 128 T; 0 U; 0 Other;  
SQ Query Match 94.7%; Score 454.4; DB 4; Length 480;  
Best Local Similarity 96.7%; Pred. No. 9e-130;  
Matches 464; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
DB 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAGATT 60  
QY 61 CTACCTCACCCTATCCATGCTTTTATAGTGAAGAATTATAGAGTATACAGCCCAATAA 120  
DB 61 TTACCTCACCCTATCCCTCTCTTAGTGAAGATTATAGAGTATACAGCCCAATAA 120  
QY 121 AAAATTGCTGCTTATAGATATACATTTTATAGAGAGCTTTTAAACGGGCAATTCCT 180  
DB 121 AAAATTGCTGCTTATAGATATACATTTTATAGAGAGCTTTTAAACGGGCAATTCCT 180  
QY 181 AATAAGCCATTTTCCGGCGTTTGTATGATAGAGGCGATGCGCAACGGGAGGTTT 240  
DB 181 AATAAGCCATTTTCCGGCGTTTGTATGATAGAGGCGATGCGCAACGGGAGGTTT 240  
QY 241 TTAGCCTTCACTAGCTTTGGGGGTTTGACCTGAAATGCCAAACAAATAATCGTGAT 300  
DB 241 TTAGCCTTCACTAGCTTTGGGGGTTTGACCTGAAATGCCAAACAAATAATCGTGAT 300  
QY 301 TTATGACGATTAAGTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 360  
DB 301 TTATGACGATTAAGTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTT 360  
QY 361 TACCATTTAGAGCTTTTAAAGTATAGGCGATCATCTGGCAAGTGGTGGCGGCTCAA 420  
DB 361 TACCATTTAGAGCTTTTAAAGTATAGGCGATCATCTGGCAAGTGGTGGCGGCTCAA 420  
QY 421 GTGGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480  
DB 421 GTGGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480

Db 121 AAAATTGCTTATAGATATACACTTTAAATGAAGATGTTTAACGGGCATTTCCCT 180  
QY 181 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAAAACGGGAGGTTT 240  
Db 181 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAAAACGGGAGGTTT 240  
QY 241 TTAGCCCTTCACTAGCTTGTGGGGGTTTGACCTCGAATCGCAAAACAAAAATCGTGTAT 300  
Db 241 TTAGCTTTTCACTAGCTTGTGGGGGTTTGACCTCGAATCGCAAAACAAAAATCGTGTAT 300  
QY 301 TTCAATGACGATTCATAGAGTTTAAATTCGGCATCCCTGTAACCCGAGCGACAGATTAGAA 360  
Db 301 TTCAATGACGATTCATAGAGTTTAAATTCGGCATCCCTGTAACCCGAGCGACAGATTAGAA 360  
QY 361 TACCATTTTGAAGTCTTTAAAGCATAAGGGCATGATCTGGCAAGTGGTGCGCGCTCAA 420  
Db 361 TACCATTTTGAAGTCTTTAAAGCATAAGGGCATGATCTGGCAAGTGGTGCGCGCTCAA 420  
QY 421 GTGGATGGCAAGTGGTGCGCTGAAGCGGAATTGAAGCCCATGATTGACAGAGAGATTAA 480  
Db 421 GTGGATGGCAAGTGGTGCGCTGAAGCGGAATTGAAGCCCATGATTGACAGAGAGATTAA 480

RESULT 3  
ABX66131  
ID ABX66131 standard; DNA; 853 BP.  
XX AC ABX66131;  
XX DT 07-MAY-2003 (first entry)  
XX DE Helicobacter pylori selected interacting domain (SID) DNA #730.  
XX DE Protein-protein interaction; ulcer; selected interacting domain; SID;  
XX KW gene, ds.  
XX OS Helicobacter pylori.  
XX PN WO200266501-A2.  
XX PD 29-AUG-2002.  
XX PF 28-DEC-2001; 2001WO-EF015428.  
XX PR 02-JAN-2001; 2001US-0259302P.  
XX PA (HYBR-) HYBRIGENICS.  
XX PA (INSP) INST PASTEUR.  
XX PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;  
XX WPI; 2002-674910/72.  
XX DR P-PSDB; ABU51387.  
XX PT New complexes of protein-protein interactions in Helicobacter pylori,  
XX PT useful for identifying modulating compounds for treating or preventing  
XX PT ulcers in mammals.  
XX PS Claim 7; Page 263; 642pp; English.  
XX CC The invention describes a complex of protein-protein interactions in  
XX CC Helicobacter pylori selected from 421 complexes given in the  
XX CC specification. The complex of protein-protein interactions are useful for  
XX CC screening for agents which modulate the interaction of proteins.  
XX CC Modulating compounds which binds to a targeted bacterial protein may be  
XX CC used for treating or preventing ulcers in a human or animal. This  
XX CC sequence encodes a selected interacting domain (SID), identified via  
XX CC protein-protein interactions  
SQ Sequence 853 BP; 258 A; 163 C; 167 G; 265 T; 0 U; 0 Other;  
Query Match 61.3%; Score 294; DB 6; Length 853;

Best Local Similarity 100.0%; Pred. No. 4.3e-80;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAACAAAGCCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60  
Db 560 ATGGAACAAAGCCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 619  
QY 61 CTACTCTACCCGCTATCCCATGCTTTAGTGAGTAGAATATAGAGTACAGCCATATAA 120  
Db 620 CTACTCTACCCGCTATCCCATGCTTTAGTGAGTAGAATATAGAGTACAGCCATATAA 679  
QY 121 AAAATTGCTTATAGATATACACTTTAAATGAAGATGTTTAACGGGCATTTCCCT 180  
Db 680 AAAATTGCTTATAGATATACACTTTAAATGAAGATGTTTAACGGGCATTTCCCT 739  
QY 181 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAAAACGGGAGGTTT 240  
Db 740 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAAAACGGGAGGTTT 799  
QY 241 TTAGCCCTTCACTAGCTTGTGGGGGTTTGACCTGAAATCGCAAAACAAAAATC 294  
Db 800 TTAGCCCTTCACTAGCTTGTGGGGGTTTGACCTGAAATCGCAAAACAAAAATC 853  
RESULT 4  
ACA30215  
ID ACA30215 standard; DNA; 441 BP.  
XX AC ACA30215;  
XX DT 19-JUN-2003 (first entry)  
XX DE Prokaryotic essential gene #11872.  
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX KW drug design; gene.  
XX OS Campylobacter jejuni.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
XX DR P-PSDB; ABU26345.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 14; SEQ ID NO 18085; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 441 BP; 152 A; 50 C; 101 G; 138 T; 0 U; 0 Other;

Query Match 28.0%; Score 134.4; DB 8; Length 441;  
 Best Local Similarity 58.4%; Pred. No. 6.6e-31;  
 Matches 256; Conservative 0; Mismatches 176; Indels 6; Gaps 1;  
 QY 39 TATAGGATATCTTACAAATCTACCTCAGCGGTATCCCATCTTTTAGTGATAGAT 98  
 DB 9 TCTATGCAATTCAGAAATCTTACCACCGTTATCTTTTACTAGTAAAT 68  
 QY 99 TATAGGATACAGCAATTAATAAATTCGCTTATAGAAATATCCTTTTAAATGAAGA 158  
 DB 69 TACAGAAATTAAGATTAAGAGATGGTGTAGGATATAAATAACATCAGTATAAGTGACCA 128  
 QY 159 CGTGTTTACGGGATTTCCCTAATAAGCCATTTTCCGGGGTTCATCGTAGAGG 218  
 DB 129 TGTATTTATGGGACATTTTCCAGGGCATCTATTTATCTGGAGTTTGTATTTAGAGG 188  
 QY 219 CATGGCGCAACCGGAGGGTTTTATAGCCTTCACTAGCTTGTGGGGTTCGACCTGAAT 278  
 DB 189 TATGCTCAACAGGTGGAGTTTACCTTTTGAAGTAT- - - - -GGAAGATAAGTGGA 242  
 QY 279 CGCCAAAACAAAATCGTGTATTTATGACGATGATAGAGTTAAATTCGCAATCCCTGT 338  
 DB 243 TCCAAAAGTAAAGTAGTTTATTTACAGGATAGATGGAGCAAAATTTAGAAATCCTGT 302  
 QY 339 AACCCAGCGGACAGATTAGATACCAATTTAGAGTCTTAAAGCATAGGCGATGATCTG 398  
 DB 303 CGCTCTGGAGATAGGCTTGTATGAAATAGCGGTGTTAAATAATCGGTGAATATGTG 362  
 QY 399 GCAAGTGGGTGCGACGGCTCAAGTGATGCGAAAGTGGTCTGTAAGCGCAATTTGAAGC 458  
 DB 363 GATTTTAAAGGCGAAGCTTTTGTAGATGGAATTTAGTTGAGAGCGCGAGCTTAAAGC 422  
 QY 459 CATGATTCAGAGAGAGA 476  
 DB 423 TATGATAGTGGATAATA 440

RESULT 5  
 ABX66187  
 ID ABX66187 standard; DNA; 659 BP.  
 XX  
 AC ABX66187;  
 XX  
 XX  
 DT 07-MAY-2003 (first entry)  
 DE Helicobacter pylori selected interacting domain (SID) DNA #786.  
 XX

KW Protein-protein interaction; ulcer; selected interacting domain; SID;  
 KW gene; db.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN WO200266501-A2.  
 XX  
 PD 29-AUG-2002.  
 XX  
 PF 28-DEC-2001; 2001WO-BP015428.  
 XX  
 PR 02-JAN-2003; 2001US-0259302P.  
 XX  
 PA (HYBR-) HYBRIGENICS.  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;  
 XX WPI; 2002-674910/72.  
 DR P-PSDB; ABU51443.  
 DR  
 XX  
 PT New complexes of protein-protein interactions in Helicobacter pylori,  
 PT useful for identifying modulating compounds for treating or preventing  
 PT ulcers in mammals.  
 XX  
 PS Claim 7; Page 277; 642pp; English.  
 CC The invention describes a complex of protein-protein interactions in  
 CC Helicobacter pylori selected from 421 complexes given in the  
 CC specification. The complex of protein-protein interactions are useful for  
 CC screening for agents which modulate the interaction of proteins.  
 CC Modulating compounds which binds to a targeted bacterial protein may be  
 CC used for treating or preventing ulcers in a human or animal. This  
 CC sequence encodes a selected interacting domain (SID), identified via  
 CC protein-protein interactions  
 XX  
 SQ Sequence 659 BP; 200 A; 125 C; 123 G; 211 T; 0 U; 0 Other;  
 Query Match 20.8%; Score 100; DB 6; Length 659;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-20;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTCAAAATT 60  
 DB 560 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTCAAAATT 619  
 QY 61 CTACCTCAGCGTATCCCATGCTTTTAGTGATAGAATTA 100  
 DB 620 CTACCTCAGCGTATCCCATGCTTTTAGTGATAGAATTA 659  
 RESULT 6  
 ADS59846  
 ID ADS59846 standard; cDNA; 459 BP.  
 XX  
 AC ADS59846;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 XX Bacterial polynucleotide #11833.  
 XX  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polynucleotide; gene; ss.  
 XX  
 OS Bacteria.  
 XX  
 XX  
 PN US2003233675-A1.  
 XX

PD 18-DEC-2003.  
 XX  
 PF 20-FEB-2003; 2003US-00369493.  
 PR  
 XX 21-FEB-2002; 2002US-0360039P.  
 PA (CAOV/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 XX  
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 XX WPI; 2004-061375/06.  
 DR  
 XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 3520; 122pp; English.  
 XX  
 CC The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polynucleotide used in  
 CC the scope of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 459 BP; 103 A; 141 C; 132 G; 83 T; 0 U; 0 Other;  
 Query Match 20.7%; Score 99.2; DB 13; Length 459;  
 Best Local Similarity 55.1%; Pred. No. 5e-20;  
 Matches 225; Conservative 0; Mismatches 168; Indels 15; Gaps 1;  
 QY 59 TTCTACTCAGCGTATCCATGCTTTTATGATGATAGATTAATAGATTAAGCCAAATA 118  
 DB 53 TCCTGCCACACCGGTATCCCTTCCTGATGATCGACGCATCGACATCGATGATGATG 112  
 QY 119 AAAAAATTGCTGCTTATAGATATACATTTTATGACGACGCTTTTACGGCATTTCC 178  
 DB 113 ATTCCGCCATCGGCATCAAGACGTCACCATCAAGCGCGCATTTTACGGCGCATTTCC 172  
 QY 179 CTAATAAGCCCATTTTCCGCGCGCTTTTCATCGTAGAGGCGCATGCGCAACCGGAGGGT 238  
 DB 173 CGGAGCAGCGGTGATGCGCGCGCTGCTGATGTCGAGCCATGCGCAGACGCGGAGCG 232  
 QY 239 TTTTACGCTTACTAGCTTGTGGGGTGTGACCTGAAATCGCAAAACAAAAATCGTGT 298  
 DB 233 CCATCTGCATCGCAGCCTCGGGCGT-----CGAAGCGCTGCTGCTGT 277  
 QY 299 ATTTCATGACGATGATAGGTTAAATTCGCATCCCTGTAAACCCAGCGGACAGATTAG 358  
 DB 278 ATTCTCTGACCATGACACGCGCAAAATTCGCAAAACCGGTGCTTCCGCGCGACAGTTGA 337

QY 359 AATACATTAGAGTCTTTAAGCATTAAGGCGCATGATCTGGCAAGTGGGTGGCACGCTC 418  
 DB 338 AGATTACGTCATAAAAAATCAAGAGCGCGCAACCTGCTCAAAATTCGCTGTGAAGCCC 397  
 QY 419 AAGTGGATGGCAAAAGTGGTCTGCTGAAGCCGAATTTGAAAGCCCATGATTG 466  
 DB 398 TGGTCGATGGACCAAGGCGCGGAGGCGGAGATTTTCAGCCCATGATGG 445  
 RESULT 7  
 ACA28687  
 ID ACA28687 standard; DNA; 432 BP.  
 XX  
 AC ACA28687;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DB Prokaryotic essential gene #10344.  
 XX  
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX  
 OS Clostridium botulinum.  
 XX  
 PN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-0299926/02.  
 DR P-PSDB; ABU24817.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 14; SEQ ID NO 16557; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway;  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product of an  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 432 BP; 169 A; 43 C; 95 G; 125 T; 0 U; 0 Other;  
Query Match 19.8%; Score 95; DB 8; Length 432;  
Best Local Similarity 53.5%; Pred. No. 9.7e-19;  
Matches 234; Conservative 0; Mismatches 185; Indels 18; Gaps 1;  
QY 34 TTTTATATAGAGCATCTTACAAATTTACCTCACCGTATCCCATGCTTTAGTGGAT 93  
DB 13 TTAGATATAAATGAATAAATAAATAATTTCCGCACAGATATCCATTTTATATAGTAGAT 72  
QY 94 AGAATTATAGAGTTACAAGCCCAATAAATAAATTTGCGCTTATAAGAAATATCACTTTTAAT 153  
DB 73 AAAATAAATCTGATTAAGAGAGGAGGAAAGCTGCGATGATATAAATAATGTACAGCTAAT 132  
QY 154 GAAGACGTTTAAACGGGCAATTCCTTAATAAGCCCAATTTTCCGGCGGCTTTGATCGTA 213  
DB 133 GAATACCTTTTAAATGGGCAATTTCCGGAAGAACCAAGTAATGCTGGGGTCTGATTATA 192  
QY 214 GAGGCGATCGCCAAACGGGAGGTTTGTAGCTTCACTAGCTTGTGGGGTTTGACCCCT 273  
DB 193 GAAGCTTTAGCACAGGTCGGAGCTGTGTCTATTTTAAAGTA-----AA 234  
QY 274 GAAATCGCCAAACAAATAATCGTGTATTTTCATGACCATTTGAAGGTATAAATCCGCATC 333  
DB 235 GAAGAGTTTAAAGGAAATAATAGCTTATTTGGAGGTATAAATAAGCTAAGTTCAGAAA 294  
QY 334 CCGTGAACCCAGGCGACAGATAGAAATACCATTTAGAAAGTCTTAAGAGCAAGGCGATG 393  
DB 295 AAGGTAGTACCAAGGAGATGTTTAAAACTTAGTATAGATCTTACTAAATAAAGCGGTT 354  
QY 394 ATCTGCAAGTGGTGGCAGGCTCAAGTGGATGGCAAGTGTCTGCAAGCGAATG 453  
DB 355 GCAGGAGTAGGTAAAGCCGTAGCTACTGTAGATGGGAAAGTAGCTGCAGAAAGCAGAATTA 414  
QY 454 AAAGCCATCATTTGCAGA 470  
DB 415 TTATTTGTAATAGAAA 431

RESULT 8  
ACA33871  
ID ACA33871 standard; DNA; 417 BP.  
XX  
AC ACA33871;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #15528.  
XX  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX  
OS Enterococcus faecium.  
XX  
PN W020027183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
DR WPI; 2003-029926/02.  
DR P-PSDB; ABU30001.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 14; SEQ ID NO 21741; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 417 BP; 147 A; 66 C; 99 G; 105 T; 0 U; 0 Other;

Query Match 19.5%; Score 93.4; DB 8; Length 417;  
Best Local Similarity 53.6%; Pred. No. 3e-18;  
Matches 230; Conservative 0; Mismatches 181; Indels 18; Gaps 1;  
QY 40 ATAGAGCATATCTTACAAATTTTACCTCACCGTATCCCATGCTTTTATGGATGAAT 99  
DB 7 ATACAGAAATATAAGAAATCATTTCCACATCGTTATCCATGCTTTTGATGATCGAGTG 66  
QY 100 ATAGAGTTACAAGCCCAATAAATAAATTTGCTGCTTATAAGAAATATCACTTTTAATGAAGAC 159  
DB 67 GAAGAAATGGTTGAAGGAGAACCGATCGTTGCTAAGAAAAATGTAAACGATCAATGAGCCC 126  
QY 160 GTGTTTAAACGGGCAATTTCCCTTAATAAGCCCAATTTTCCGGGCGTTTTCATCTAGAGGCG 219  
DB 127 TTTTTCAGAGACATTTTCCAGAGAAACCTGTTATGCCAGGGGTATGATGTTGAAGCA 186  
QY 220 ATGGCGAAACCGGAGGGTTTTATAGCCCTTCACTAGCTTTGTGGGGTGTGACCCCTGAAATC 279  
DB 187 ATGGCTCAAGCAGAGCAGTGTGCACTTCTTCT-----CTTGAACAA 228  
QY 280 GCCAAAAACAAAATCGGTGTTATTTTCATGACGATTAAGTGTAAATTCGCGATCCCTGTA 339  
DB 229 TTCAAGGAAAAGCGGCTTATTTTCGCGGATTTGGACAAAGCAAAATTCGTAAGAAGTT 288



QY 340 ACCCAGCGCAGAGATTAGATACCACTTTAGAGTCTTAAAGCATTAAGGCGCATGCTGG 399  
 |||||  
 Db 289 ACACCGGGAGATACCTTATACCTTGAAGTTGAGATATTGAAGTGAAAGCTTCAGCCGGA 348  
 |||||  
 QY 400 CAAAGTGGTGGCAGCGCTCAAGTGAATGCGCAAAAGTGGTCTGCTGAAGCCGGAATTGAAAGCC 459  
 |||||  
 Db 349 ATAGGCAAGGAATCGCTAAGTAGACGGTAAAGAGTAGCCGAGCGGAATTAACCTTT 408  
 |||||  
 QY 460 ATGATTGCA 468  
 |||||  
 Db 409 ATGATTGGA 417  
 |||||

RESULT 9  
 ACA53362  
 ID ACA53362 standard; DNA; 462 BP.  
 AC ACA53362;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Prokaryotic essential gene #35019.  
 XX  
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX  
 OS Vibrio cholerae.  
 XX  
 PN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI: 2003-029926/02.  
 DR P-PSDB; ABU49492.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 41232; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway; (8) required for proliferation, or that inhibits cellular proliferation; (9) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of the strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 462 BP; 120 A; 95 C; 112 G; 135 T; 0 U; 0 Other;

Query Match 19.3%; Score 92.6; DB 8; Length 462;

Best Local Similarity 53.8%; Pred. No. 5.5e-18;

Matches 222; Conservative 0; Mismatches 179; Indels 12; Gaps 1;

QY 59 TTCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAAGCAATA 118  
 |||||  
 Db 56 TTCTGCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAAGCAATA 115  
 |||||  
 QY 119 AAAAAATTGCGCTTATAAGAAATATCACTTTTAATGAAGACGTGTTTAACGGGCAATTCC 178  
 |||||  
 Db 116 AGTACCTGATTGGTTTGA AAAATGTTTTCAGTGAATGAGCCTCAGTTTACGGGTCACTTTC 175  
 |||||  
 QY 179 CTAATAAGCCCATTTTCCCGGGCTTTTTCGATGAGGGCATGCGCAAAACGGGAGGCT 238  
 |||||  
 Db 176 CACAACCTGCGGGTTTCCCGGGTGTATTGATTTTAGAAGCCATGGCTCAAGCCACTGGTC 235  
 |||||  
 QY 239 TTTTAGCTTTCACCTAGCTTGTGGGGTTTGACCTGAAATCGCCAAACAAAAATCGTGT 298  
 |||||  
 Db 236 TGTGGGTTCAAAACCTTTGTGGCCAA-----AGAGATGACTGTACT 283  
 |||||  
 QY 299 ATTTCATGAGCATGTGATAAGGTTAAATTCGCGCATCCCTGTAAACCCAGGCGACAGATTAG 358  
 |||||  
 Db 284 ATTTGCCAGCATGTGATAAGCTAAGTTCCGTAAGCAGTGACACCGGGCGATCAACTGA 343  
 |||||  
 QY 359 AATACCATTTAGAGTCTTAAAGCATTAAGGCATGATCTGCAAGTGGTGGCAGGCTC 418  
 |||||  
 Db 344 TGGTTGAAGTTGAATTCCTTAAAGAGCGTCTGTGATCGCGCTGTTTAAACGGCGTGGCA 403  
 |||||  
 QY 419 AAGTGTAGTGC AAAAGTGTGCTGTAAGCCGAATTTGAAAGCCATGATTGCAGAG 471  
 |||||  
 Db 404 AAGTGTAGTGTGATGTGCTCTCTTCTGCTCACTGAATAATGGCAGCTCGAGAG 456  
 |||||

RESULT 10

ACA36484

ID ACA36484 standard; DNA; 435 BP.

AC ACA36484;

XX

DT 19-JUN-2003 (first entry)

XX

DE Prokaryotic essential gene #18141.

XX

KW Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

XX

OS Listeria monocytogenes.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.



PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zykkind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU32614.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PI for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 24354; 1766pp; English.  
PS  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway;  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 435 BP; 142 A; 61 C; 104 G; 128 T; 0 U; 0 Other;

RESULT 11

ABQ69109  
ID ABQ69109 standard; DNA; 1755 BP.  
XX  
XX AC ABQ69109;  
XX  
XX DT 29-AUG-2003 (revised)  
DT 29-AUG-2002 (first entry)  
XX  
XX DE *Listeria monocytogenes* 4b contig DNA sequence #1875.  
XX  
XX KW Antibacterial; *Listeria*; food contamination; mutational analysis;  
KW infection; ds.  
XX  
XX OS *Listeria monocytogenes* ATCC 19115.  
XX  
XX PN WO200228891-A2.  
XX  
XX PD 11-APR-2002.  
XX  
XX PF 04-OCT-2001; 2001WO-FR003061.  
XX  
XX PR 04-OCT-2000; 2000FR-00012697.  
XX  
XX PA (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
PI Kunst F, Glaser P;  
XX  
XX WPI; 2002-332479/37.  
XX  
XX PT New genomic sequences from *Listeria* species, useful for detection,  
PT treatment and prevention of infection, also related polypeptides,  
PT antibodies and modulators.  
XX  
XX PS Claim 14; SEQ ID NO 1922; 180pp; French.  
XX  
XX CC The present invention relates to nucleic acid sequences (ABQ67188-  
CC ABQ71212) from *Listeria* sp. The sequences are useful as probes and  
CC primers for identification and/or detection of *Listeria* (e.g. as  
CC contaminants in foods, or mutational analysis) and for analysis of gene  
CC expression. Proteins encoded by the nucleic acid sequences can be used to  
CC screen for compounds that modulate gene expression, replication and  
CC pathogenicity of *Listeria* (potential therapeutic agents), also for  
CC treating infections by *Listeria*, and are useful as immunogens in anti-  
CC *Listeria* vaccines. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. (Updated  
CC on 29-AUG-2003 to standardise OS field)  
XX  
XX SQ Sequence 1755 BP; 583 A; 238 C; 385 G; 548 T; 0 U; 1 Other;

Query Match 18.5%; Score 88.6; DB 6; Length 1755;  
Best Local Similarity 64.3%; Pred. No. 1.7e-16;  
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 34 TTTTATAGACATATCTTACAAATTCCTACCGCTATCCCATGCTTTTAGTGAT 93  
DB 830 TTAGATATTAAGAAATCAAGAGATTTTGCCTCATCTTATCCATTTTGTAGTAG 889  
QY 94 AGAATTATAGAGTTACAAGCCAATAAAAAATGTGCTTATAAGAAATATCACTTTAAT 153  
DB 890 AGAGTTATTTCTGTGTAAGGAAGGCAAAAAGTTACCGCTATTAAAGATGTAACAGCAAT 949  
QY 154 GAAGACGTGTTTAACGGCAATTTCCCTAATAAGCCATTTTCCCGGGGCTTTTGATCGTA 213  
DB 950 GAAGAATTTTTTAACGGGCACCTTTCTCGAGTATCTCTGTAATGCCAGGAGTTTAAATAGTG 1009  
QY 214 GAGGCGATGGCGCAAAACGGAGGGTTT 240  
DB 1010 GAAGCTCTAGCGCAAACTAGTGGAAAT 1036

RESULT 12

ABQ70710  
ID ABQ70710 standard; DNA; 1927 BP.  
XX AC ABQ70710;  
XX XX  
XX 29-AUG-2003 (revised)  
XX 29-AUG-2002 (first entry)  
XX XX  
XX Listeria monocytogenes 4b contig DNA sequence #652.  
XX DE  
XX Antibacterial; Listeria; food contamination; mutational analysis;  
XX KW infection; ds.  
XX XX  
XX Listeria monocytogenes ATCC 19115.  
XX OS  
XX WO200228891-A2.  
XX PN  
XX XX  
XX PD 11-APR-2002.  
XX XX  
XX PF 04-OCT-2001; 2001WO-FR003061.  
XX PR  
XX 04-OCT-2000; 2000FR-00012697.  
XX XX  
XX PA (INSP ) INST PASTEUR.  
XX PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX PI  
XX Kunst F, Glaser P;  
XX XX  
XX WPI; 2002-332479/37.  
XX DR  
XX XX  
XX PT New genomic sequences from Listeria species, useful for detection,  
XX PT treatment and prevention of infection, also related polypeptides,  
XX PT antibodies and modulators.  
XX XX  
XX Claim 14; SEQ ID NO 3523; 180pp; French.  
XX PS  
XX CC The present invention relates to nucleic acid sequences (ABQ67188-  
XX CC ABQ71212) from Listeria sp. The sequences are useful as probes and  
XX CC primers for identification and/or detection of Listeria (e.g. as  
XX CC contaminants in foods, or mutational analysis) and for analysis of gene  
XX CC expression. Proteins encoded by the nucleic acid sequences can be used to  
XX CC screen for compounds that modulate gene expression, replication and  
XX CC pathogenicity of Listeria (potential therapeutic agents), also for  
XX CC treating infections by Listeria, and are useful as immunogens in anti-  
XX CC Listeria vaccines. Note: The sequence data for this patent did not form  
XX CC part of the printed specification, but was obtained in electronic format  
XX CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. (Updated  
XX CC on 29-AUG-2003 to standardise OS field)  
XX XX  
XX SQ Sequence 1927 BP; 641 A; 264 C; 426 G; 596 T; 0 U; 0 Other;  
  
Query Match 18.5%; Score 88.6; DB 6; Length 1927;  
Best Local Similarity 64.3%; Pred. No. 1.8e-16;  
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
  
QY 34 TTTTATATAGCATATCTTACAAATTCACCTCACCGCTATCCCATGCTTTTAGTGAT 93  
Db 990 TTAGATATTAAAGAAATCAAGAGATTTTGCCTCATCGTTATCCATTTTGTAGTAGAT 1049  
  
QY 94 AGAATTATAGAGTTACACCAATAAAAAATGTCGCTTATAAGAAATATCACTTTTAAT 153  
Db 1050 AGAGTTATTCTGTTGAAGAGCAAAAAAGTTACCGCTATTAAAGAAATGTAACAGCAAT 1109.  
  
QY 154 GAAGACGTTTAAACGGGCAATTCCTTAATAGCCCAATTTCCCGGGCGTTTGTATCGTA 213  
Db 1110 GAAGAATTTTTTAAACGGGCACCTTCTCGTAGTATCCTGTAATGCCAGGAGTTTTAATAGTG 1169  
  
QY 214 GAGGCGATCGCGCAACCGGAGGTTT 240  
Db 1170 GAAGCTCTAGCGCAAACTAGTGAATT 1196  
  
RESULT 13

ABA03041\_25/c  
Continuation (26 of 30) of ABA03041 from base 2500001 (Listeria monocytogenes EGD-e genom  
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041  
WP Fragment Name Begin End  
WP ABA03041\_00 1 110000  
WP ABA03041\_01 100001 210000  
WP ABA03041\_02 200001 310000  
WP ABA03041\_03 300001 410000  
WP ABA03041\_04 400001 510000  
WP ABA03041\_05 500001 610000  
WP ABA03041\_06 600001 710000  
WP ABA03041\_07 700001 810000  
WP ABA03041\_08 800001 910000  
WP ABA03041\_09 900001 1010000  
WP ABA03041\_10 1000001 1110000  
WP ABA03041\_11 1100001 1210000  
WP ABA03041\_12 1200001 1310000  
WP ABA03041\_13 1300001 1410000  
WP ABA03041\_14 1400001 1510000  
WP ABA03041\_15 1500001 1610000  
WP ABA03041\_16 1600001 1710000  
WP ABA03041\_17 1700001 1810000  
WP ABA03041\_18 1800001 1910000  
WP ABA03041\_19 1900001 2010000  
WP ABA03041\_20 2000001 2110000  
WP ABA03041\_21 2100001 2210000  
WP ABA03041\_22 2200001 2310000  
WP ABA03041\_23 2300001 2410000  
WP ABA03041\_24 2400001 2510000  
WP ABA03041\_25 2500001 2610000  
WP ABA03041\_26 2600001 2710000  
WP ABA03041\_27 2700001 2810000  
WP ABA03041\_28 2800001 2910000  
WP ABA03041\_29 2900001 2944528  
  
Query Match 18.5%; Score 88.6; DB 6; Length 110000;  
Best Local Similarity 64.3%; Pred. No. 1.1e-15;  
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
  
QY 34 TTTTATATAGCATATCTTACAAATTCACCTCACCGCTATCCCATGCTTTTAGTGAT 93  
Db 102680 TTAGATATTAAAGAAATCAAGAGATTTTGCCTCATCGTTATCCATTTTGTAGTAGAT 102621  
  
QY 94 AGAATTATAGAGTTACACCAATAAAAAATGTCGCTTATAAGAAATATCACTTTTAAT 153  
Db 102620 AGAGTTATTCTGTTGAAGAGCAAAAAAGTTACCGCTATTAAAGAAATGTAACAGCAAT 102561  
  
QY 154 GAAGACGTTTAAACGGGCAATTCCTTAATAGCCCAATTTCCCGGGCGTTTGTATCGTA 213  
Db 102560 GAAGAATTTTTTAAACGGGCACCTTCTCGTAGTATCCTGTAATGCCAGGAGTTAATAGTG 102501  
  
QY 214 GAGGCGATCGCGCAACCGGAGGTTT 240  
Db 102500 GAAGCTCTGCGCAAAACAGTGAATT 102474  
  
RESULT 14  
ABA03041\_26/c  
Continuation (27 of 30) of ABA03041 from base 2600001 (Listeria monocytogenes EGD-e genom  
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041  
WP Fragment Name Begin End  
WP ABA03041\_01 100001 210000  
WP ABA03041\_02 200001 310000  
WP ABA03041\_03 300001 410000  
WP ABA03041\_04 400001 510000  
WP ABA03041\_05 500001 610000  
WP ABA03041\_06 600001 710000  
WP ABA03041\_07 700001 810000  
WP ABA03041\_08 800001 910000  
WP ABA03041\_09 900001 1010000  
WP ABA03041\_10 1000001 1110000  
WP ABA03041\_11 1100001 1210000



Search completed: January 12, 2006, 01:29:30  
Job time : 464 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:07:09 ; Search time 3071 Seconds

(without alignments)  
7312.862 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggacaagcatcaaaaa.....tgattgcagagagatttaa 480

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	85	17.7	724	6	CD830511	CD830511 BN40.045M
2	84.8	17.7	851	8	DR926513	DR926513 EST111805
3	84.4	17.6	565	5	BQ577106	BQ577106 PESToabl
4	84	17.5	845	4	BN8092WL	BN8092WL Arabidops
5	83.6	17.4	939	9	BZ567344	BZ567344 pac92-164
6	83.2	17.3	820	8	DR948072	DR948072 EST113961
7	83.2	17.3	830	8	DR942398	DR942398 EST113393
8	83.2	17.3	841	8	DR938067	DR938067 EST112960
9	83.2	17.3	848	8	DR929905	DR929905 EST112144
10	83.2	17.3	877	8	DR931821	DR931821 EST112336
11	83	17.3	465	1	AJ750880	AJ750880 AJ750880
12	82.8	17.2	631	8	DN910931	DN910931 Cuc6_3 Cl
13	82.4	17.2	583	1	AI994492	AI994492 70149847
14	82.4	17.2	598	1	AU235591	AU235591 AU235591
15	82.4	17.2	774	4	CNS0A1U5	BN834110 Arabidops
16	82.4	17.2	799	4	CNS0929N	BN833007 Arabidops
17	82.4	17.2	840	4	CNS0A1T2	BN833404 Arabidops
18	82	17.1	848	8	DR938068	DR938068 EST112960
19	81.8	17.0	764	7	CNT734884	CNT734884 25RDNM_U
20	81.6	17.0	693	8	CK045132	CK045132 UCRCS07_2
21	80.8	16.8	457	5	BQ791261	BQ791261 E4137 Ch1
22	80.8	16.8	471	2	BG544298	BG544298 E2138 Ch1

23	80.6	16.8	538	7	CK937166	CK937166 CGF100451
24	80.6	16.8	728	8	CX542609	CX542609 UCRPT01_5
c 25	80.6	16.8	791	8	CV887178	CV887178 UCRCS04_2
c 26	80.6	16.8	845	7	CF832673	CF832673 UCRCS02_0
27	80.2	16.7	616	6	CD822254	CD822254 BN25.0441
28	80.2	16.7	819	8	DR948073	DR948073 EST113961
29	80.2	16.7	852	8	DR929906	DR929906 EST112144
30	79.8	16.6	854	8	DR925940	DR925940 EST111747
c 31	79.6	16.6	629	6	CF515461	CF515461 CAP0001_1
c 32	79.6	16.6	703	6	CB971995	CB971995 CAB10006
c 33	79.6	16.6	747	6	CB972347	CB972347 CAB10006
c 34	79.6	16.6	760	6	CB972067	CB972067 CAB10006
c 35	79.6	16.6	766	6	CB972267	CB972267 CAB10006
c 36	79.6	16.6	772	6	CF210321	CF210321 CAB20006
c 37	79.6	16.6	786	5	BQ796669	BQ796669 EST 5607
38	79.6	16.6	870	6	CF515372	CF515372 CAP0001_1
39	79.6	16.6	885	6	CF210235	CF210235 CAB20006
40	79.2	16.5	547	8	DN776664	DN776664 E4179 CG
41	79.2	16.5	586	3	BP530330	BP530330 BP530330
42	79	16.5	515	7	CO908062	CO908062 BJ0202650
43	79	16.5	539	8	CV882970	CV882970 Mdfit3116
44	79	16.5	570	8	CV881159	CV881159 Mdfit3113
45	79	16.5	683	7	CN181917	CN181917 UCRCS04_0

#### ALIGNMENTS

RESULT 1  
LOCUS CD830511 724 bp mRNA linear EST 10-JUL-2003  
DEFINITION BN40.045M05F011228 BN40 Brassica napus cDNA clone BN40045M05, mRNA sequence.  
ACCESSION CD830511  
VERSION CD830511.1 GI:32512451  
KEYWORDS EST.  
SOURCE Brassica napus (rape)  
ORGANISM Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 724)  
AUTHORS Genoplatte.  
TITLE Genoplatte, a major partnership french program in plant genomics  
JOURNAL Unpublished (2003)  
COMMENT Contact: Genoplatte  
Genoplatte  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>)  
and <http://genoplatte-info.inbioigen.fr>.

#### FEATURES

source Location/Qualifiers  
1..724  
/organism="Brassica napus"  
/mol\_type="mRNA"  
/cultivar="Jet Neuf"  
/db\_xref="taxon:3708"  
/clone="BN40045M05"  
/tissue\_type="seed"  
/clone\_lib="BN40"

#### ORIGIN

Query Match 17.7%; Score 85; DB 6; Length 724;  
Best Local Similarity 64.5%; Pred. No. 1.3e-13;  
Matches 127; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 40 ATAGACATCTTACAAATCTTACCTCCCGTATCCCATGCTTTTAGTGATAGTAATT 99  
DB 303 ATTAACAGATAGAGAGATTTGCGTTCACAGGTTCCCGTTCTGTTAGTGATAGTAGTG 362  
QY 100 ATAGAGTTACAGCCCAATAAAAAAATTGCGCTTATAAGAAATATCACTTTTAAATGAAGAC 159

Db 363 ATAGAGTACACAGCTGGTGAATCTCGGCTAGCTATCAAGAAGCTTACCATTATGCAAT 422  
QY 160 GTGTTTAAACGGGCAATTCCTTAATAGCCCAATTTCCCGGGCGTTTGTGATCGTAGAGGGC 219  
Db 423 TTCTTTCCCTGGGCAATTTCCCGAGAGGCCCATATGCTTGGTCTCATGTTGAGGCC 482  
QY 220 ATGGCGCAACCGGAGG 236  
Db 483 ATGGCTCAGTGGGAGG 499

RESULT 2  
LOCUS DR926513/c  
DEFINITION EST1118052 Aquilegia cDNA library Aquilegia formosa x Aquilegia pubescens cDNA clone COINE67, mRNA sequence.  
ACCESSION DR926513  
VERSION DR926513.1 GI:71695876  
KEYWORDS EST.  
SOURCE Aquilegia formosa x Aquilegia pubescens  
ORGANISM Aquilegia formosa x Aquilegia pubescens  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales; Ranunculaceae; Aquilegia.  
REFERENCE 1 (bases 1 to 851)  
AUTHORS Hodges,S.A., Rensink,W., Buell,C.R., Borevitz,J., Kramer,E., Nordborg,M. and Tomkins,J.  
Generation of ESTs from Aquilegia  
Other ESTs: EST1118053  
Contact: Scott Hodges  
Department of Ecology, Evolution and Marine Biology  
University of California, Santa Barbara  
Santa Barbara, CA 93106, USA  
Tel: 805 893 7813  
Fax: 805 893 4724  
Email: hodges@lifesci.ucsb.edu  
Seq primer: TTTTCTTTTCTTTTCTTTT (where N = A, G & C).

FEATURES  
source  
1. .851  
/organism="Aquilegia formosa x Aquilegia pubescens"  
/mol\_type="mRNA"  
/db\_xref="taxon:338618"  
/clone="COINE67"  
/tissue\_type="mixed shoot and floral apical meristems, flower buds, leaves and roots"  
/lab\_host="DH10B T1 (T1 and T5 phage resistance)"  
/clone\_lib="Aquilegia cDNA library"  
/notes="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI; F2, F3, and F4 lines of Aquilegia formosa X A. pubescens were grown from seed in greenhouses at UC Santa Barbara. From these plants three sets of tissue were collected: 1) Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively), 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN  
Query Match 17.7%; Score 84.8; DB 8; Length 851;  
Best Local Similarity 50.6%; Pred. No. 1.5e-13;  
Matches 243; Conservative 0; Mismatches 222; Indels 15; Gaps 1;

QY 1 ATGGAAACAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
Db 640 ATTGAATAAAGATTTCAAAGCGTTTCCAACAATAATGGATATGAAAAGATTCGTAACATT 581  
QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGATTTAGAGTTACAGGCAATAAA 120  
Db 580 TTGCCCTCATAGGTTTCGTTTCTTAGTTGATAGAGTGAATGATATGATATGATCCTGGAGTT 521  
QY 121 AAAATTGTGCTTTATAGAATATCACTTTTAAAGACGCTGTTTAAACGGGCAATTCCTT 180  
Db 520 TCGGCGTTTGCATTAAAGATGTTACGATTAAATGATAATTTCTTCCAGGCCAATTTTCTCT 461  
QY 181 AATAAGCCCAATTTCCCGGGCGTTTTCATCGTAGAGGCAATGGCGCAACGGGAGGTTT 240  
Db 460 GAGAGGCCAATTTATGCTGCTGTTTCTTATGATTGAGGCAATGGCTCAAGTTGGAGGCTTG 401  
QY 241 TTAGCCTTCACTAGCTTGTGCGGGTTTGACCTCGAATCGCCAAACAAAATCGTGAT 300  
Db 400 GTAATGCTGACCCAGAAAGTGGGTGTTCTCGTG-----AAAATTTCTTC 356  
QY 301 TTCTATGACCAATTGATAAGGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 360  
Db 355 TTTCAGGAATTCACAAGTCAGATTCGCGAAGCTGTGGTTGCGAGGGATACGCTTGTA 296  
QY 361 TACCATTTAGAGTCTTTAAAGCATAGAGGCGATGATCTGGCAAGTGGGTGGACGCTCAA 420  
Db 295 ATGAGAATGATTCTCATAAAGAAACGCTTTGGAGTAGTTAAGATGGAGGGGAAGGCTTAT 236  
QY 421 GTGGATGGCAAGTGGTCTGCGTGAAGCGCAATTTGAAGCCATGATTGCGAGAGAGATTAA 480  
Db 235 GTTGGGGGTGAGGTGGTATGTGACGGAGAGTTACGATGGCTCTTTGGGAAGAGCAATGA 176

RESULT 3  
LOCUS BQ577106  
DEFINITION PFESToab13h01.y1 Plasmidium falciparum 3D7 asexual cDNA Plasmidium falciparum 3D7 CDNA 5' similar to TR:O97376 O97376  
BETA-HYDROXYACYL-ACP DEHYDRATASE PRECURSOR. i, mRNA sequence.  
ACCESSION BQ577106  
VERSION BQ577106.1 GI:21480423  
KEYWORDS EST.  
SOURCE Plasmidium falciparum 3D7  
ORGANISM Plasmidium falciparum 3D7  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.  
REFERENCE 1 (bases 1 to 565)  
AUTHORS Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,F., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I., Teagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A., Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D., Waterston,R., Wilson,R. and Sibley,D.  
WashU Plasmidium EST Project  
Unpublished (2001)  
CONTACT: L. David Sibley  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Library was constructed by Debopam Chakrabarti DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: L. David Sibley (sibley@borcim.wustl.edu), Washington University  
Seq primer: -40UP from Gibco  
High quality sequence stop: 423.

FEATURES  
source  
1. .565  
/organism="Plasmidium falciparum 3D7"  
/mol\_type="mRNA"  
/db\_xref="taxon:36329"  
/lab\_host="DH10B (GeneHog, Invitrogen, Inc.)"





source 1...939  
/organism="Pseudomonas aeruginosa"  
/mol\_type="Genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pac2-164\_6892"  
/clone\_lib="pac2-164"  
/note="clinical isolate 2-164 Whole genomic shotgun library."

## ORIGIN

Query Match 17.4%; Score 83.6; DB 9; Length 939;  
Best Local Similarity 53.9%; Pred. No. 3.5e-13;  
Matches 223; Conservative 0; Mismatches 179; Indels 12; Gaps 2;  
QY 44 AGCATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGATATATAG 103  
DB 65 AGCTTGATATCAATCTCTGACGCCCGCTATCCATCTCTGCTGGATCGGTGTAG 124  
QY 104 AGTTCAAGCCAA---TAAAAAAATGTGCTTTATGAAGATATCACTTTTAATGAAGACG 160  
DB 125 AGCTGACATCGAGGCAAGCGCATTCGCGCTTACAAGATGTCAATGAGCCGT 184  
QY 161 TGTTTAAACGGGATTTCCCTATATAGCCATTTTCCGGCGGTTTGTATCGTAGAGGCA 220  
DB 185 TCTTCAACGGACATTTCCCGGAGCACCCGATCATGCGGGCGGTGCTGATCATCGAGGCGA 244  
QY 221 TGGCGCAACAGGGGGGTTTATAGCTTCACTAGCTTGTGGGGTTTGACCCCTGAAATCG 280  
DB 245 TGGCCAGCGCGCGGTATCTCGGTTTCA-----AGATGCTCGACGTGAAGCCCG 295  
QY 281 CCAAAACAAAATCTGTGATTTTATGACGATGATGAAGTTAAATTCGCGATCCCTGTAA 340  
DB 296 CGACGGCACCTTTACTTCTGCTGCGTTCGACAGCTGCGCTTCCGCCAGCGGTC 355  
QY 341 CCCAGGCGACAGATPAGATACCATTTAGAAGTCTTAAAGCATAAGGCGATGATCTGGC 400  
DB 356 TGCAGGCGGACCAACTGCACTGACCGCAAGTTTCACTGAGCGTGAAGCGCAGCATCTGGA 415  
QY 401 AGTGGTGGCAGCGCTCAAGTGGATGGAGTGTGCTGCTGAAGCGGAATGA 454  
DB 416 AATTCGACTGCCACGCTACCGTGCAGCAGCAAGCCGGTATGCTCGGCTGAAATCA 469

## RESULT 6

DR948072/c 820 bp mRNA linear EST 02-AUG-2005  
LOCUS  
DEFINITION  
EST1139611 Aquilegia cDNA library Aquilegia formosa x Aquilegia  
pubescens cDNA clone CO1R896, mRNA sequence.

DR948072

DR948072.1 GI:71717435

EST.

Aquilegia formosa x Aquilegia pubescens

Aquilegia formosa x Aquilegia pubescens

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;

Ranunculaceae; Aquilegia.

1 (bases 1 to 820)

Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,

Nordborg, M. and Tomkins, J.

Generation of ESTs from Aquilegia

Unpublished (2005)

Other ESTs: EST1139612

Contact: Scott Hodges

Department of Ecology, Evolution and Marine Biology

University of California, Santa Barbara

Santa Barbara, CA 93106, USA

Tel: 805 893 7813

Fax: 805 893 4724

Email: hodges@lifesci.ucsb.edu

Seq primer: TTTTCTTTTCTTTTCTTTTCTTTT (where N = A, G &amp; C).

Location/Qualifiers

1..820

## FEATURES

source

/organism="Aquilegia formosa x Aquilegia pubescens"  
/mol\_type="mRNA"  
/db\_xref="taxon:338618"  
/clone="CO1R896"  
/tissue\_type="mixed shoot and floral apical meristems,  
flower buds, leaves and roots"  
/lab\_host="DH10B T1 (T1 and T5 phage resistance)"  
/clone\_lib="Aquilegia cDNA library"  
/note="Vector: pCMV SPORT6.1; Site 1: EcoRI, Site 2: NotI;  
F2, F3, and F4 lines of Aquilegia formosa x A. pubescens  
were grown from seed in greenhouses at UC Santa Barbara.  
From these plants three sets of tissue were collected: 1)  
Small flower buds (<10 mm) and very young inflorescences  
(T1 & 29% by weight respectively), 2) Medium (7-20 mm) and  
large (at or near anthesis) flower buds (65 & 35% by  
weight respectively) and 3) Shoot apical meristems. A  
fourth set of tissue was collected from plants of A.  
formosa. These plants were grown from seed in sand and at  
approximately 1 month root tissue and leaf tissue of  
various developmental stages were collected (84 & 16% by  
weight respectively). Total RNA was extracted from each  
set of tissue and pooled in the following proportions:  
1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled  
total RNA, mRNA was extracted and enriched for full-length  
messages and then normalized with proprietary methods by  
Invitrogen."

## ORIGIN

Query Match 17.3%; Score 83.2; DB 8; Length 820;  
Best Local Similarity 50.4%; Pred. No. 4.4e-13;  
Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;  
QY 1 ATGGAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTCAAAATT 60  
DB 641 ATTGAAAAGATTCAAAGCGTTTCCAACTAATGATATGAAAGATTCGTAACATT 582  
QY 61 CTACCTCACCGCTATCCATGCTTTTAGTGGATAGATATAGAGTTACAGGCAATAA 120  
DB 581 TTGCCTCATAGGTTTCGGTTTCTTCTAGTTGATAGAGTGAATGAATATGATCTCGAGTT 522  
QY 121 AATTTGTGCTTATAGATATACCTTTTAAAGACAGCTTTTAAACGGCATTTCCCT 180  
DB 521 TCAGCGGTTGCTATTAAGAAATGTACGATTAATGAATATTTCTTCCAGGCGCATTTTCT 462  
QY 181 AATAAGCCCATTTTCCCGGCGTTTTCATCGTAGAGGCGATGCGCAACCGGAGGTTT 240  
DB 461 GAGAGGCAATTAATGCTGCTGTTCTTATGATGAGGCAATGCTCAAGTTGGAGGCTTG 402  
QY 241 TTAGCCTTCACTAGCTTTGCGGGGTTTGAACCTGAATTCGCCAAAACAAAATCGTGTAT 300  
DB 401 GTAAATGCTGACCCCGAGAAAGTGGGTGTTCTCGTG-----AAAATTTCTTC 357  
QY 301 TTATGACGATTTGATAGGTTAAATTCGCTATCCGCTATCCCTTAACCCCGGCGACAGATTAGAA 360  
DB 356 TTATCAGGAATTCGCAAGTCAGATTCGCGAAACCTGTTGTTGCGAGGAGTACCGCTTGA 297  
QY 361 TACCATTTAGAAGTCTTAAAGCATAAGGCGCATGATCTGGCAAGTGGGTGGCCACGCTCAA 420  
DB 296 ATGAGATGATTTCTATTAAGAAACCGCTTTGGAGTAGTTAGATGAGGGGAGGCGCTAT 237  
QY 421 GTGGATGGCAAGTGGTTCGCTGAAGCCGAAATGAAAGCCATGATTCGACAGAGAGATTAA 480  
DB 236 GTTGGGGTGAAGTGGTATGTGACGGAGAGTTTCCAGATGGCTCTTTGGGAAGCGGAATGA 177

## RESULT 7

DR942398/c 830 bp mRNA linear EST 02-AUG-2005  
LOCUS  
DEFINITION  
EST1133937 Aquilegia cDNA library Aquilegia formosa x Aquilegia  
pubescens cDNA clone CO1Q471, mRNA sequence.  
ACCESSION  
DR942398  
DR942398.1 GI:71711761  
KEYWORDS  
EST.



SOURCE Aquilegia formosa x Aquilegia pubescens  
 ORGANISM Aquilegia formosa x Aquilegia pubescens  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;  
 Ranunculaceae; Aquilegia.  
 REFERENCE 1 (bases 1 to 830)  
 AUTHORS Hodges S.A., Rensink W., Buell C.R., Borevitz J., Kramer E.,  
 Nordborg M. and Tomkins J.  
 TITLE Generation of ESTs from Aquilegia  
 JOURNAL Unpublished (2005)  
 COMMENT Other ESTs: EST1133938  
 Contact: Scott Hodges  
 Department of Ecology, Evolution and Marine Biology  
 University of California, Santa Barbara  
 Santa Barbara, CA 93106, USA  
 Tel: 805 893 7813  
 Fax: 805 893 4724  
 Email: hodges@lifesci.ucsb.edu  
 Seq primer: TTTT TTTT TTTT TTTT TTTT TTTT N (where N = A, G & C).  
 FEATURES  
 source  
 1..830  
 Location/Qualifiers  
 /organism="Aquilegia formosa x Aquilegia pubescens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:338618"  
 /clone="COL0A71"  
 /tissue\_type="mixed shoot and floral apical meristems,  
 flower buds, leaves and roots"  
 /lab\_host="DH10B T1 (T1 and T5 phage resistance)"  
 /clone\_lib="Aquilegia cDNA library"  
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens  
 were grown from seed in greenhouses at UC Santa Barbara.  
 From these plants three sets of tissue were collected: 1)  
 Small flower buds (<10 mm) and very young inflorescences  
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and  
 large (at or near anthesis) flower buds (65 & 35% by  
 weight respectively) and 3) Shoot apical meristems. A  
 fourth set of tissue was collected from plants of A.  
 formosa. These plants were grown from seed in sand and at  
 approximately 1 month root tissue and leaf tissue of  
 various developmental stages were collected (84 & 16% by  
 weight respectively). Total RNA was extracted from each  
 set of tissue and pooled in the following proportions:  
 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled  
 total RNA, mRNA was extracted and enriched for full-length  
 messages and then normalized with proprietary methods by  
 Invitrogen."  
 ORIGIN  
 Query Match 17.3%; Score 83.2; DB 8; Length 830;  
 Best Local Similarity 50.4%; Pred. No. 4.5e-13;  
 Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;  
 QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGATATCTTACAAATT 60  
 DB 641 ATTGAAAAAGATTCAAAGCGTTTCCAACTAATGATATTGAAAGATTCTGTAACATT 582  
 QY 61 CTACTTACCGCTATCCCATGCTTTTAGTGGATAGATATATAGATTACAAACCAATATA 120  
 DB 581 TTGCGCTCATAGTTTCCGTTCTTCTAGTTGATAGATGATTAATGATATGATCTGGAGTT 522  
 QY 121 AAAATTGTCGCTTAAAGAAATATCACTTTTAAATGAAGAGCTGTTTAAACGGGCAATTCCTT 180  
 DB 521 TCGGGGGTGTCTATTAGATAGTTAGATTAATGATTAATTTCTTCCAGGCCATTTCTCT 462  
 QY 181 AATAAGCCCAATTTTCCGGGGCTTTTGTATCGTAGAGGGCATGCGCAACACGGGAGGTTT 240  
 DB 461 GAGAGGCCAATTATGCTGGTGTCTTATGATTGAGGCAATGCTCAAGTTGGAGGCTTG 402  
 QY 241 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTGAAATCGCCAAACAAAATTCGTGTAT 300  
 DB 401 GTAATGCTGACCCCAAGAGTGGGTGTTCTCGTG-----AAAATTTCTTC 357

QY 301 TTCAATGACGATTGATAAGGTTAAATTCGGCATCTCCTGTAAACCCAGGCGACAGATTAGAA 360  
 DB 356 TTTACAGGAATTGACAAAGTCAGATTCCGGAACCTGTGGTTCAGGGGATACGCTTTGTA 297  
 QY 361 TACCATTTAGAAGTCTTTAAAGCATAAAGGGCATGATCTGCCAAGTGGTGGCAGCGCTCAA 420  
 DB 296 ATGAGATGATTTCTCATTAAGAAACGCTTTGGAGTAGTTAAGATGAGGGGAAGCGCTAT 237  
 QY 421 GTGATGCGCAAGTGGTGGCTGAAGCCGAATTGAAGCCATGATTGCGAGAGAGATTAA 480  
 DB 236 GTTGGGGTGGAGTGGTATGTGACGAGAGTTCCAGATGGCTCTTGGGAAGAGCGAATGA 177  
 RESULT 8  
 DR938067/c  
 LOCUS  
 DEFINITION EST1129606 Aquilegia cDNA library Aquilegia formosa x Aquilegia  
 pubescens cDNA clone COLPL21, mRNA sequence.  
 ACCESSION DR938067  
 VERSION DR938067.1 GI:71707430  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM  
 Aquilegia formosa x Aquilegia pubescens  
 Aquilegia formosa x Aquilegia pubescens  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;  
 Ranunculaceae; Aquilegia.  
 REFERENCE 1 (bases 1 to 841)  
 AUTHORS Hodges S.A., Rensink W., Buell C.R., Borevitz J., Kramer E.,  
 Nordborg M. and Tomkins J.  
 TITLE Generation of ESTs from Aquilegia  
 JOURNAL Unpublished (2005)  
 COMMENT Other ESTs: EST1129607  
 Contact: Scott Hodges  
 Department of Ecology, Evolution and Marine Biology  
 University of California, Santa Barbara  
 Santa Barbara, CA 93106, USA  
 Tel: 805 893 7813  
 Fax: 805 893 4724  
 Email: hodges@lifesci.ucsb.edu  
 Seq primer: TTTT TTTT TTTT TTTT TTTT TTTT N (where N = A, G & C).  
 FEATURES  
 source  
 1..841  
 Location/Qualifiers  
 /organism="Aquilegia formosa x Aquilegia pubescens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:338618"  
 /clone="COLPL21"  
 /tissue\_type="mixed shoot and floral apical meristems,  
 flower buds, leaves and roots"  
 /lab\_host="DH10B T1 (T1 and T5 phage resistance)"  
 /clone\_lib="Aquilegia cDNA library"  
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens  
 were grown from seed in greenhouses at UC Santa Barbara.  
 From these plants three sets of tissue were collected: 1)  
 Small flower buds (<10 mm) and very young inflorescences  
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and  
 large (at or near anthesis) flower buds (65 & 35% by  
 weight respectively) and 3) Shoot apical meristems. A  
 fourth set of tissue was collected from plants of A.  
 formosa. These plants were grown from seed in sand and at  
 approximately 1 month root tissue and leaf tissue of  
 various developmental stages were collected (84 & 16% by  
 weight respectively). Total RNA was extracted from each  
 set of tissue and pooled in the following proportions:  
 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled  
 total RNA, mRNA was extracted and enriched for full-length  
 messages and then normalized with proprietary methods by  
 Invitrogen."  
 ORIGIN  
 Query Match 17.3%; Score 83.2; DB 8; Length 841;  
 Best Local Similarity 50.4%; Pred. No. 4.5e-13;  
 Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;



COMMENT Other\_ESTs: EST1123361  
Contact: Scott Hodges  
Department of Ecology, Evolution and Marine Biology  
University of California, Santa Barbara  
Santa Barbara, CA 93106, USA  
Tel: 805 893 7813  
Fax: 805 893 4724  
Email: hodges@lifesci.ucsb.edu  
Seq primer: TTTTCTTTTCTTTTCTTTTCTTTT (where N = A, G & C).  
Location/Qualifiers  
source  
1. .877  
/organism="Aquilegia formosa x Aquilegia pubescens"  
/mol\_type="mRNA"  
/db\_xref="taxon:338618"  
/clone="CO10876"  
/tissue\_type="mixed shoot and floral apical meristems,  
flower buds, leaves and roots"  
/lab\_host="DH10B T1 (T1 and T5 phage resistance)"  
/clone\_lib="Aquilegia cDNA library"  
/notes="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;  
F2, F3, and F4 lines of Aquilegia formosa X A. pubescens  
were grown from seed in greenhouses at UC Santa Barbara.  
From these plants three sets of tissue were collected: 1)  
Small flower buds (<10 mm) and very young inflorescences  
(T1 & 29% by weight respectively), 2) Medium (7-20 mm) and  
large (at or near anthesis) flower buds (65 & 35% by  
weight respectively) and 3) Shoot apical meristems. A  
fourth set of tissue was collected from plants of A.  
formosa. These plants were grown from seed in sand and at  
approximately 1 month root tissue and leaf tissue of  
various developmental stages were collected (84 & 16% by  
weight respectively). Total RNA was extracted from each  
set of tissue and pooled in the following proportions:  
1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled  
total RNA, mRNA was extracted and enriched for full-length  
messages and then normalized with proprietary methods by  
Invitrogen."

FEATURES  
source  
1. .877  
/organism="Gerbera hybrid cv. 'Terra Regina',  
Paulin, L., Ainaoja, M., Kotilainen, M., Koskela, S., Teeri, T.H. and  
Elomaa, P."  
/tissue\_type="mRNA"  
/db\_xref="taxon:226891"  
/clone="G0000100013G01F1"  
/tissue\_type="late petal"  
/dev\_stage="5-9"  
/clone\_lib="G00001"

ORIGIN  
Query Match 17.3%; Score 83.2; DB 8; Length 877;  
Best Local Similarity 50.4%; Pred. No. 4.5e-13;  
Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;  
QY 1 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60  
DB 641 ATTGAABAGATTCAAGGGTTTCCACAATATGATATGAAGATTCTGTAACATT 582  
QY 61 CTACCTCAGCGCTATCCCATGCTTTTATAGAGCATATCTTACAAAT 120  
DB 581 TTGCCTCATAGGTTTCCGTTTCTTCTAGTTGATAGATGATGAATATGATCTCGAGTT 522  
QY 121 AAAATGTGCTTATAGAAATATCACTTTTAAAGACGTTTAAACGGGCAATTTCCCT 180  
DB 521 TCGGGGTTGCTATTAAGAAATGTACGATTAATGATAATTTCTTCCAGGCCAATTTTCC 462  
QY 181 AATAAGCCCAATTTCCCGGCGTTTTCATCGTAGAGGCGATGCGCAACGGGAGGTTT 240  
DB 461 GAGAGCCAAATTATGCTGCTGTTCTTATGATTTGAGCAATGCTCAAGTTGAGGCTTG 402  
QY 241 TTAGCCCTTCACTAGCTTTGTTGGGGTTTGAACCTTGAATCGCCAAAACAAAATCGTGTAT 300  
DB 401 GTAATGCTGACCCAGAGTGGGTGTTCTCGTG-----AAAATTTCTTC 357  
QY 301 TTCATGACAAATTGATAGGTTAAATTCGGCATCCCTGTATACCCAGCGCAGATAGAA 360  
DB 356 TTACAGGAAATTGACAAAGTCAGATTCCGGAACCTGTGTTGCGAGGGATACGTTGTA 297  
QY 361 TACCATTAGAGTCTTAAAGCATAGGCGATGATCTGCGCAAGTGTGCGCGCTCAA 420  
DB 296 ATGAGAAATGTTCTCAATTAAGAAACGCTTTGGAGTAGTTAAGATGGGGGAAGGCGCTAT 237  
QY 421 GTGGATGGCAAGTGTGCTGCTGAAGCGCAATTTGAAGGCCATGATTCAGAGAGAGATTAA 480

Db 236 GTTGGGGTGAGTGGTATGTGACGGAGATTACGATGGCTCTTGGGAAGACGAATGA 177

RESULT 11  
AJ750880  
LOCUS  
DEFINITION  
AJ750880 G00001 Gerbera hybrid cv. 'Terra Regina' cDNA clone  
G000100013G01F1, mRNA sequence.  
AJ750880  
VERSION  
AJ750880.1 GI:58321534  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Gerbera hybrid cv. 'Terra Regina',  
Paulin, L., Ainaoja, M., Kotilainen, M., Koskela, S., Teeri, T.H. and  
Elomaa, P.  
TITLE  
Analysis of the floral transcriptome uncovers new regulators of  
organ determination and gene families related to flower organ  
differentiation in Gerbera hybrida (Asteraceae)  
JOURNAL  
Genome Res. 15 (4), 475-486 (2005)  
PUBMED  
15781570  
COMMENT  
Contact: Alatalo ER  
Institute of Biotechnology  
Viikinkaari 9, P.O. Box 56, 00014 University of Helsinki, FINLAND.  
Location/Qualifiers  
source  
1. .465  
/organism="Gerbera hybrid cv. 'Terra Regina',  
Paulin, L., Ainaoja, M., Kotilainen, M., Koskela, S., Teeri, T.H. and  
Elomaa, P."  
/mol\_type="mRNA"  
/cultivar="Terra Regina"  
/db\_xref="taxon:226891"  
/clone="G000100013G01F1"  
/tissue\_type="late petal"  
/dev\_stage="5-9"  
/clone\_lib="G00001"

ORIGIN  
Query Match 17.3%; Score 83; DB 1; Length 465;  
Best Local Similarity 58.8%; Pred. No. 4.5e-13;  
Matches 143; Conservative 0; Mismatches 100; Indels 0; Gaps 0;  
QY 1 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60  
DB 153 ATTGAATCAGGTTCCCTCGGTTTCCAAACCGTTCTTGTATATCAATCACAATCGGACATT 212  
QY 61 CTACCTCAGCGCTATCCCATGCTTTTATAGTATGATAGAAATATAGAGTTACAAGCCAATAA 120  
DB 213 TTACCACACGATTTCCATTCTTTGGTGGATAGGTTGATGATCAATCCAGGTGTT 272  
QY 121 AAAATGTGCTTATAGAAATATCACTTTTAAAGACGTTTAAACGGGCAATTTCCCT 180  
DB 273 TCAGCAGTTGCGATCAAGAATGTGACTATCAATGATACTTCTTCTCGGCGATTTCTCT 332  
QY 181 AATAAGCCCAATTTCCCGGCGTTTTCATCGTAGAGGCGATGCGCAACGGGAGGTTT 240  
DB 333 GAGAGGCGGATTAAGCTGCTGTTCTTATGTCGAGCAATGCGCGCAATGCTGCTGTTG 392  
QY 241 TTA 243  
DB 393 GTA 395

RESULT 12  
DN910931/c  
LOCUS  
DEFINITION  
DN910931 Cuc6\_3C12 Cucumber fruit development cDNA library Cucumis sativus  
cDNA, mRNA sequence.  
DN910931  
ACCESSION  
DN910931.1 GI:62895694

**KEYWORDS**  
SOURCE  
ORGANISM

Cucumis sativus (cucumber)  
Cucumis sativus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.

**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**  
**COMMENT**

Vergara, G., McGrath, J.M. and Grumet, R.  
Cucumber fruit cDNA library  
Unpublished (2005)  
Contact: Rebecca Grumet  
Rebecca Grumet  
Michigan State University  
Horticulture Department, Michigan State University, East Lansing,  
MI 48824, USA  
Tel: 517 353 0890  
Fax: 517 355 5191 x431  
Email: grumet@msu.edu  
Plate: Cuck\_3 row: C column: 12.

**FEATURES**  
source  
1..631  
/organism="Cucumis sativus"  
/mol\_type="mRNA"  
/db\_xref="taxon:3659"  
/tissue\_type="fruit"  
/dev\_stage="2 days to 50 days post-pollination"  
/lab\_host="E. coli XL1-Blue MRP"  
/clone\_lib="Cucumber fruit development cDNA library"  
/note="Organ: fruit; Vector: UNI-ZAP; cDNAs were made from  
a total of 10 grams of combined tissues from developing  
pickling cucumber fruits. Fruits were harvested and  
finely ground in liquid nitrogen every two days after  
pollination for 50 days. Tissue was combined within an  
interval in equal amounts, and combined in the following  
interval amounts: 2 to 6 days contributed 1.5 grams of the  
total combined 10 grams (15%); 8 to 14 days (40%); 16 to  
18 days (5%); 22 to 28 days (10%); 30 to 36 days (10%); 38  
to 44 days (10%); and 46 to 50 days (10%)."

**ORIGIN**  
Query Match 17.2%; Score 82.8; DB 8; Length 631;  
Best Local Similarity 63.6%; Pred. No. 5.e-13;  
Matches 126; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 39 TATAGACATCTTACAAATTTACTCCCTATCCCTGCTTTTAGTGATAGAT 98  
|||  
DB 622 TATCAACGAGATTCGTAAAATTTGGCTCACAGGTTTCGATCTTTTAGTAGATAGAT 563  
|||  
QY 99 TATAGAGTTACAAGCCAATAAAAAATTGCGCTTATAAGAATATCATCTTTTAATGAAGA 158  
|||  
DB 562 GATGGATACATCTCTGGAGTCTCGGTTGTGCTATTAGATGTCACCTAATAAGAGAG 503  
|||  
QY 159 CGTGTTTAACGGGCATTTCCCTAATAAGCCCAATTTTCCCGGGCGTTTTCATCGTAGAGG 218  
|||  
DB 502 TTCTTTTCCAGGGCACTTCCCTGAGAGACCACATCATGCTGGTGTCTCATGCTTGAGGC 443  
|||  
QY 219 CATGGCGCAAACGGGAGG 236  
|||  
DB 442 AATGGCTCAACATGGCGG 425  
|||

**RESULT 13**  
AI994492  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AI994492 583 bp mRNA linear EST 08-SEP-1999  
701498447 A. thaliana, Ohio State clone set Arabidopsis thaliana  
cDNA clone 701498447, mRNA sequence.

AI994492 AI994492.1 GI:5841397  
EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**  
**COMMENT**

Chen, J., Hillman, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,  
Wang, X., Mamiya, K., Kim, C., Doyle, M., Brzoska, P.,  
Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D.,  
Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C.,  
Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,  
Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and  
Hanson, D.  
Arabidopsis thaliana Gene Expression MicroArray  
Unpublished (1999)  
Contact: David Smoller, Ph.D.  
Genome Systems, Inc., a wholly owned subsidiary of Incyte  
Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: service@genomesystems.com.

**FEATURES**  
source  
1..583  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone\_lib="A. thaliana, Ohio State clone set"  
/note="cDNA library was made from selected clones from the  
Arabidopsis thaliana Ohio State clone set."

**ORIGIN**  
Query Match 17.2%; Score 82.4; DB 1; Length 583;  
Best Local Similarity 59.3%; Pred. No. 7.1e-13;  
Matches 140; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCAATCTCAATTTTTATAGAGCATATCTTACAAATT 60  
|||  
DB 263 ATCGAATAGATGATGAGCGTTTCGACAGTGATGACATCAATAGATACAGAGATT 322  
|||  
QY 61 CTACCTCACAGATTCCTTTCTGTTAGTGTAGATGATGATGATGATGATGATGATGATGAT 120  
|||  
DB 323 TTGCTCTACAGATTCCTTTCTGTTAGTGTAGATGATGATGATGATGATGATGATGATGAT 382  
|||  
QY 121 AAAATTGCTTTAAGAATATCACTTTTAAAGAAGAGCTGTTTAAACGGGCAATTTCCCT 180  
|||  
DB 383 TCTCGGTAGCTTATTAATAAAGCTTACCATTAATGCAATTTCTTCTCGGCAATTTCCCT 442  
|||  
QY 181 ATAAGCCCAATTTCCCGGCCGTTTGTATCGTAGGGCATGCGCAACGGGAGG 236  
|||  
DB 443 GAGAGACCTTAATGCTGGAGTTCTCATGTTGAGGCAATGCTCAAGTGGGAGG 498  
|||

**RESULT 14**  
AU235591 598 bp mRNA linear EST 01-APR-2002  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AU235591 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-24-C08 5',  
mRNA sequence.

AU235591 AU235591.1 GI:19874760  
EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**  
**COMMENT**

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,  
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,  
and Shinozaki, K.  
Large scale analysis of Arabidopsis full-length CDNA (2002b)  
Unpublished (2002)  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group

Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr))  
Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:14:29 ; Search time 153 Seconds  
(without alignments)  
5576.664 Million cell updates/sec

Title: US-10-662-126-36  
Perfect score: 480  
Sequence: 1 atggaacaagcatcaaaaa.....tgattgcagagagattaa 480

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/PCUS COMB.seq.\*  
7: /cgn2\_6/prodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/prodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/prodata/1/ina/baCkfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	18.0	429	3	US-09-134-000C-1148
2	82.8	17.2	465	3	US-09-252-991A-4232
3	82.8	17.2	483	3	US-09-252-991A-4647
4	81.8	17.0	489	3	US-09-543-681A-2455
5	78.2	16.3	561	3	US-09-489-039A-5222
6	77.2	16.1	423	3	US-09-583-110-1245
7	77.2	16.1	441	3	US-09-107-433-289
8	76.6	16.0	558	3	US-09-540-236-1553
9	76.6	16.0	99229	3	US-09-596-002-37
10	75.6	15.8	423	3	US-09-196-388-1
11	75.6	15.8	483	3	US-10-089-019-17
12	75.6	15.8	19702	3	US-08-961-527-7
13	74.2	15.5	10303	3	US-09-634-238-410
14	73.4	15.3	462	3	US-09-902-540-3820
15	73.4	15.3	21010	3	US-09-902-540-1188
16	72.2	15.0	1830121	3	US-09-557-884-1
17	72.2	15.0	1830121	3	US-09-643-990A-1
18	72.2	15.0	1830121	3	US-10-158-865-1
19	69.6	14.5	461	3	US-09-339-614-1
20	69.6	14.5	501	3	US-10-089-019-9
21	69.2	14.4	471	3	US-09-134-001C-1397
22	65.4	13.6	1029	3	US-09-221-017B-988
23	62.2	13.0	492	3	US-09-328-352-3116
24	60.8	12.7	2926	3	US-09-710-279-3747

25	60.8	12.7	6309	3	US-08-956-171E-480	Sequence 480, App
26	60.8	12.7	6309	3	US-08-781-986A-480	Sequence 480, App
27	59.6	12.4	354	3	US-09-107-532A-1534	Sequence 1534, App
c 28	58.4	12.2	1230025	3	US-09-198-452A-1	Sequence 1, Appli
c 29	58.4	12.2	1230230	3	US-09-438-185A-1	Sequence 1, Appli
30	56.6	11.8	471	3	US-09-902-540-4150	Sequence 4150, Ap
c 31	56.6	11.8	18469	3	US-09-902-540-1205	Sequence 1205, Ap
32	44	9.2	591	3	US-09-543-681A-2590	Sequence 2590, Ap
33	35.6	7.4	822	3	US-09-710-279-727	Sequence 727, App
34	35.6	7.4	3215	3	US-09-710-279-3566	Sequence 3566, Ap
c 35	35.6	7.4	3993	3	US-09-710-279-3985	Sequence 3985, Ap
36	35	7.3	288031	3	US-09-949-016-14864	Sequence 14864, A
37	34.6	7.2	640681	3	US-09-790-988-1	Sequence 1, Appli
c 38	34.4	7.2	601	3	US-09-949-002-9724	Sequence 9724, App
c 39	34.4	7.2	36677	3	US-09-949-002-832	Sequence 832, App
c 40	34.2	7.1	7218	2	US-08-232-463-14	Sequence 14, Appl
c 41	34	7.1	412	3	US-09-270-767-3889	Sequence 3889, Ap
c 42	34	7.1	412	3	US-09-270-767-19171	Sequence 19171, A
c 43	34	7.1	83178	3	US-09-949-016-14606	Sequence 14606, A
c 44	33.8	7.0	969	3	US-09-252-991A-4308	Sequence 4308, Ap
c 45	33.6	7.0	16995	3	US-08-961-527-82	Sequence 82, Appl

## ALIGNMENTS

### RESULT 1

US-09-134-000C-1148  
; Sequence 1148, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1148  
; LENGTH: 429  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis  
; US-09-134-000C-1148

Query Match		18.0%;	Score 86.4;	DB 3;	Length 429;
Best Local Similarity		53.6%;	Pred. No. 5.4e-18;		
Matches		236;	Conservative	0;	Mismatches 186; Indels 18; Gaps 2;
Qy	32	AAATTTTATAGACATATCTTACAAATTTACCTCACCGCTATCCCATGCTTTTAGTGG	91		
Db	8	AAATTAACAATTTACAGAAATTTACAGAAATTTTACCATCGCTATCCCTTTTATTATTAG	67		
Qy	92	ATAGAATTTATAGAGTCAACAGCAATAAAAAATTTGCTTATAAGAAATATCACTTTTA	151		
Db	68	ATAGTGTGTGAGAAGTCAATCCCGGGGACGCGTTGTAGCGAAAGAAATGTAACGGTTA	127		
Qy	152	ATGAAGACGTGTTTAAACGGGCAATTTCCCTAATAAGCCATTTCCCGGCGTTTGTATCG	211		
Db	128	ATGAGCAAGTTTTTCAAGGTCATTTCTCGGAATCTGTTTTCGACGAGTTTTTAATTA	187		
Qy	212	TAGAGGCATGGCGCAACCGGAGGTTTTTAGCCTTCACTAGCTTGTGGGGGTTTGACC	271		
Db	188	TTGAATCGTTGGCTCAAGCAGGCGCGCTGGCATTTTCAA-----TGC	232		
Qy	272	CTGAATTCGCAACAAACAAATCTGTATTTCATGACGATTGATAAGGTAAATTCGCA	331		
Db	233	CTGAATTT---CAAGGGAACCGGCATCTTTGGTGAATTAGTAAGCAAAATTTAGAC	289		
Qy	332	TCCTGTAAACCCAGCGCAGACGATTTAGAAATCCATTTTAGAAGTCTTAAAGCATAGGGCA	391		



Db 290 AAAAGGTAACACCGGACACATTAATTTTAGAGGTTGAACCTTTTAAAGATCCGGCTT 349  
Qy 392 TGATCTGGCAAGTGGGTGCGACGGCTCAAGTGGATGGCAAGTGGTTCGCTGAAGCCGAAT 451  
Db 350 CTGCTGGAATGGTAAAGGTGTGGCAAAAGTTAAATGGAATAAAGTTGCGGAGGCTGAAT 409  
Qy 452 TCAGAGCCATGATTCGAG 471  
Db 410 TAACTTTATGATGGATAG 429

RESULT 2  
US-09-252-991A-4232/c  
; Sequence 4232, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4232  
; LENGTH: 465  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4232

Query Match 17.2%; Score 82.8; DB 3; Length 465;  
Best Local Similarity 54.2%; Pred. No. 8.4e-17;  
Matches 218; Conservative 0; Mismatches 172; Indels 12; Gaps 2;

Qy 56 AAATCTACTCACCCTATCCCATGCTTTTAGTGATAGAATTATAGAGTTTACAAGCCA 115  
Db 410 AATACCTGCTCATCGTATCTTCTGCTGTGTGATCGGTGTAGAGTGCATCG 351  
Qy 116 A---TAAAAAATGTGCTTATAAGAAATACATTTTAAATGAAGACGTGTTTAAACGGGC 172  
Db 350 AGGCAAGCGCATTCGCGCTACAAGATGTCAATATCAATGAGCGGTCTTCAACGGAC 291  
Qy 173 ATTTCCCTTAATAGCCATTTCCCGGCGTTTGTATCGTAGAGGCATGGCCAAACGG 232  
Db 290 ATTTCCCGAGCACCGCATGTCGCGGCGTGTGTATCATCGAGGCGATGGCCAGCGG 231  
Qy 233 GAGGGTTTTAGCTTCACTAGCTTTGTGGGGTTTGACCTGAAATCGCAAAACAAAA 292  
Db 230 CCGGTATCTCGTTTCA-----AGATGCTCGAGTGAAGCCGCGACGCCACCC 180  
Qy 293 TCGTGTAATTCATGACGATGATGAAGTTTAAATTCGCGATCCCTGTAACCCAGGCGACA 352  
Db 179 TTTACTACTTGTGCTGTTCCGCAAGCTGCGCTTCCGCGAGCGGTCTGCGGGGAC 120  
Qy 353 GATTGAATACCATTTAGAAGTCTTAAAGCATGAAGGCATGATCTGGCAAGTGGTGGA 412  
Db 119 AATGCAACTGCACGCAAGTTTCATCAGCGTGAAGCGCAGCATCTGGAATTCGACTGCC 60  
Qy 413 CGGCTCAAGTGGTGAAGTGGTTCGCTGAAGCCGAATTGA 454  
Db 59 ACGTACCGTGCAGCAGCAAGCCGGTATGCTCGGCTGAAATCA 18

RESULT 3  
US-09-252-991A-4647  
; Sequence 4647, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4647  
; LENGTH: 483  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4647

Query Match 17.2%; Score 82.8; DB 3; Length 483;  
Best Local Similarity 54.2%; Pred. No. 8.6e-17;  
Matches 218; Conservative 0; Mismatches 172; Indels 12; Gaps 2;

Qy 56 AAATCTACTCACCCTATCCCATGCTTTTAGTGGATAGAATTATAGATTACAAGCCA 115  
Db 68 AATACCTGCTCATCGCTATCTTCTGCTGTGTGATCGGTGTAGAGCTGGACATCG 127  
Qy 116 A---TAAAAAATGTGCTTATAAGAAATACATTTTAAATGAAGACGTGTTTAAACGGGC 172  
Db 128 AGGCAAGCGCATTCGCGCTACAAGAAATGTCAATATCAATGAGCCGTCTTCAACGGAC 187  
Qy 173 ATTTCCCTTAATAGCCATTTCCCGGCGTTTGTATCGTAGAGGCATGGCCAAACGG 232  
Db 188 ATTTCCCGAGCACCGCATGTCGCGGCGTGTGTATCATCGAGGCGATGGCCAGCGG 247  
Qy 233 GAGGGTTTTAGCTTCACTAGCTTTGTGGGGTTTGACCTGAAATCGCAAAACAAAA 292  
Db 248 CCGGTATCTCGTTTCA-----AGATGCTCGAGTGAAGCCGCGACGCCACCC 298  
Qy 293 TCGTGTAATTCATGACGATGATGAAGTTTAAATTCGCGATCCCTGTAACCCAGGCGACA 352  
Db 299 TTTACTACTTGTGCTGTTCCGCAAGCTGCGCTTCCGCGACGCGTCTCTCGGGGCGACC 358  
Qy 353 GATTGAATACCATTTAGAAGTCTTAAAGCATGAAGGCATGATCTGCAAGTGGGTGGCA 412  
Db 359 AACTGCAACTGCAGCCCAAGTTTCATCAGCGTGAAGCGCAGCATCTGGAATTCGACTGCC 418  
Qy 413 CGGCTCAAGTGGATGGCAAAAGTGGTTCGCTGAAGCCGAATTGA 454  
Db 419 ACGTACCGTGCAGCAGCAAGCCGGTATGCTCGGCTGAAATCA 460

RESULT 4  
US-09-543-681A-2455  
; Sequence 2455, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 2455  
; LENGTH: 489  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-2455

Query Match 17.0%; Score 81.8; DB 3; Length 489;  
Best Local Similarity 52.0%; Pred. No. 1.8e-16;  
Matches 216; Conservative 0; Mismatches 187; Indels 12; Gaps 1;

Qy 40 ATAGAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAATT 99



Db 61 ATCGAAGAAATTTAGATTCTTCCACCGTTACCAATTTTATTGTTGACCGTGC 120  
QY 100 ATAGAGTTACAGCCATAAAAAATTTGTCGCTTATAAGAAATATCACTTTTAATGAAGAC 159  
Db 121 CTTCGATTTTGAAGAAGGAAATTTTAAAGACAGTAAAAAATGTATCTTTTAATGAGCCT 180  
QY 160 GTGTTTAAACGGGATTTCCCTAATAAGCCATTTTCCGGGCGTTTATGATCGTAGAGGC 219  
Db 181 TTCTTTCAAGTCACATTTCCAGCAAACTTAATTTCCCTGGTGTACTGCTTCTTGAAGCA 240  
QY 220 ATGCGCAAAACGGGAGGTTTTTAGCTTCACTAGCTTGTGGGGGTTTGACCCCTGAAATC 279  
Db 241 ATGCGCGACGACAGATTTTATGATTCAAAAGCTAGAAAACTGGAAACCG----- 295  
QY 280 GCCAAACAAAAATCGTGTATTTTCATGACGATTTGAAGGTTAAATTCGCATCCCTGTA 339  
Db 296 -----GTGAACCTTTACTATTTTGCTGCTATTGACGGTCTCGCTTTAAAGCTCCTGTC 348  
QY 340 ACCCGAGGCGACAGNTAGATACCATTTAGAAGCTTTAAAGCATTAAGGGCATGATCGG 399  
Db 349 TTACCTGGAGACCAAAATGGTTCTTGAAGTTGAATTCATTAAGAAACGCCCGTGGCTTGA 408  
QY 400 CRAAGTGGTGGCAGCGCTCAAGTGGATGCAAGTGTGCTGCTGAAGCGCAATTGA 454  
Db 409 CGCTTTAAAGGCGCTGCTAAGTTGATGAGAAATTTGCCTGCGAAGCTGAGATGA 463

RESULT 5  
US-09-489-039A-5222  
; Sequence 5222, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 5222  
; LENGTH: 561  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5222

Query Match 16.3%; Score 78.2; DB 3; Length 561;  
Best Local Similarity 58.3%; Pred. No. 3e-15;  
Matches 137; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
QY 40 ATAGAGCATATCTTACAAATTTACTCTACCGCTATCCCATGCTTTTAGTGGATAGATT 99  
Db 133 ATGAAAGAAATTTGGAACCTTCTGCTCACCGTTACCGCTTCTGCTGGTAGACCGCGTG 192  
QY 100 ATAGAGTTACAAGCCAAATAAAAAATTTGTCGCTTATAAGAAATATCACTTTTAATGAAGAC 159  
Db 193 CTGGATTTTGAAGAGGTGCTTTCTGCGCGCATGAAAAATGTTCTGTAAACAGACCG 252  
QY 160 GTGTTTAAACGGGATTTCCCTAATAAGCCATTTTCCGGGCGTTTATGATCGTAGAGGC 219  
Db 253 TTTTCCAGGGGCACTTCCCTGTGAAGCGATTTTCCGGGCGTCTGATTTCTGGAAGCA 312  
QY 220 ATGCGCAAAACGGGAGGTTTTTAGCTTCACTAGCTTGTGGGGGTTTGACCCCTG 274  
Db 313 ATGCGCGAAGCAACCGGATTTGCGGTTTCAAAAGCGTCCGAAAACTGGAACACAG 367

RESULT 6  
US-09-583-110-1245  
; Sequence 1245, Application US/09583110  
; Patent No. 6699703

; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 1245  
; LENGTH: 423  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-1245

Query Match 16.1%; Score 77.2; DB 3; Length 423;  
Best Local Similarity 55.5%; Pred. No. 5.4e-15;  
Matches 223; Conservative 0; Mismatches 158; Indels 21; Gaps 3;  
QY 39 TATAGACATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAT 98  
Db 9 TATTCAAGGAATCAAGAAGCCCTTCCCAACCGTTATCTCTATGCTTCTAGTGGACCGTGT 68  
QY 99 TATAGAGTTACAGCCCAATAAAAAATTTGCGCTTATAAGAAATATCACTTTTAATGAAGA 158  
Db 69 CTGGAGGT---GAGCGAGGATACCATTTGCTTATCAAAAAATGTACCATCAACGAGCC 125  
QY 159 CGTGTTTTAAACGGGCAATTTCCCTAATAAGCCCATTTTCCCGGCGTTTGTAGTCGTAGAGGG 218  
Db 126 TTCTTTTAAACGGCCATTTCTCTCAATACCCAGTTATGCCAGGTGTTCTGATATGAAGC 185  
QY 219 CATGGCGCAAAACGGGAGGTTTTTAGCTTCTACAGCTTGTGGGGGTTTGACCCCTGAAT 278  
Db 186 CTGGCGCAAACTGCCGCTGTTGGAGTTATCAA-----AACCTGAAA- 229  
QY 279 CGCCAAACCAAAATCGTGTATTTTCATGACGATTTAGAGTTAAATTCGCCATCCCTGT 338  
Db 230 --ATAAGGAAAACTGCTCTTTTACGCTGGTATGATAGGTTAAGTTCAAGAAGCAAGT 287  
QY 339 AACCCGAGCGCACAGATTAGATACCATTTTAGAAGTCTTAAAGCATTAAGGCGCATGCTG 398  
Db 288 TGTACCAAGCGGACCAATTTGTTATGACAGCGACTTTTGTAAACGTCGTGGCACCATAGC 347  
QY 399 GCAAGTGGTGGCAGCGCTCAAGTGGATGGCAAGTGGTGGC 440  
Db 348 TGTGGTTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 389

RESULT 7  
US-09-107-433-289  
; Sequence 289, Application US/09107433  
; Patent No. 6800744  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID  
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE  
; THERAPEUTICS  
; NUMBER OF SEQUENCES: 5206  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: <Unknown>

RESULT 8  
US-09-540-236-1553  
; Sequence 1553, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.

```

RESULT 9
US-09-596-002-37
; Sequence 37, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 99629
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 78467
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte template ID No. 6632636 37
; PUBLICATION INFORMATION:
US-09-596-002-37

```

Query Match 16.0%; Score 76.6; DB 3; Length 99629;  
Best Local Similarity 51.9%; Pred. No. 1.4e-13;  
Matches 203; Conservative 0; Mismatches 179; Indels 9; Gaps 1;

QY 44 AGCATATCTTACAAATCTACCTCACCGCTATCCATGCTTTTAGTGAGATATATAG 103  
DB 69789 ACCAAATTAAAGCATATTTGCCGATCGCTATCCATCATGCTGATTCAGCGAGTACCG 69848

QY 104 AGTTACAAGCCAATAAAAAATTTGCTTTATAGAAATATCACTTTTAAATGAAGACGTGT 163  
DB 69849 CTTGTAAACCAATGATGATGATCACAGGATATAAAAAATATCAGTATTAATGAAGAGTTAT 69908

QY 164 TTAACGGGCAATTTCCCTAATAAGCCATTTTCCGGGCGTTTGTATCGTAGAGGCGATGG 223  
DB 69909 TTAACGGGTCAATTTTCCAGATAATCAATCATGCGAGCGCTACTACAAGTTGAAGCAATGG 69968

QY 224 CGCAAAACGGGAGGTTTATAGCTTCACATGCTTTGTTGGGGTTTGACCCGAAATCGCCA 283  
DB 69969 CACAGCTATCGGTATTTTAGTGTTCATCATGCGAGGCGAGCGCTGATG----- 70019

QY 284 AAACAAAAATCGTGTATTTTCATGACGATTTAGATTAAGTTAAATTCGCGATCCCTGTAAACCC 343  
DB 70020 ATGTTATCTGTACTTGTGTCAGGTGTGATAAAGTTCTGTTTAAAAAAGTCTTACCT 70079

QY 344 CAGGCGACAGATTAGAAATACCATTTAGAGTCTTAAAGCATAAAGGCGATGATCTGGCAAG 403  
DB 70080 CAGGCGATCAGCTGTTTATTCGCTCAAGGTTTGTATGAATAAACGAGATATTTATAAAT 70139

QY 404 TGGGTGCGACGCTCAAGTGGATGCGCAAGT 434  
DB 70140 TTGAGTGTACCGCCCATGTTGATGGTCAATT 70170

RESULT 10  
US-09-196-388-1  
; Sequence 1, Application US/09196388  
; Patent No. 6277595  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Patrick  
; APPLICANT: Konstantinidis, Alex  
; APPLICANT: Russell, Robert  
; TITLE OF INVENTION: fab2  
; FILE REFERENCE: GM10182  
; CURRENT APPLICATION NUMBER: US/09/196,388  
; CURRENT FILING DATE: 1998-11-19  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 423  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(420)  
US-09-196-388-1

Query Match 15.8%; Score 75.6; DB 3; Length 423;  
Best Local Similarity 55.2%; Pred. No. 1.8e-14;  
Matches 222; Conservative 0; Mismatches 159; Indels 21; Gaps 3;

QY 39 TATAGAGCATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAT 98  
DB 9 TATTCAGGAATCAAGAGAGCTCTTCCCAACCGTTATCTATGCTTCTAGTGGACCGTGT 68

QY 99 TATAGAGTTACAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAAATGAAGA 158  
DB 69 CTTTGA---AGTGAGGAGGATACCATTTGCTATCAAAATGTGACCATCAACGAGCC 125

QY 159 CGTGTTTAACGGGCATTTCCCTAATAAGCCATTTTCCCGGCGTTTGTAGTCGTAGAGGG 218  
DB 129 CTTTGA---AGTGAGGAGGATACCATTTGCTATCAAAATGTGACCATCAACGAGCC 185

QY 186 TTTCTTTAACGGGCATTTCTCTCAATACCCAGTTATGCCAGTGTGTGTGATTATGGAAGC 245  
DB 219 CATGGCGCAAAACGGGAGGTTTGTAGCTTCACTAGCTTTGGGGGTTTGACCCCTGAAT 278

QY 246 CTTTGGCGCAAACTGCCGCTGTGTGGAGTTATCAA-----AACCTGAAA- 289

QY 279 CGCCAAACCAAAATCGTGTATTTTCATGACGATTTAGAGTCTTAAAGCATAAAGGCGCATGAT 338  
DB 290 --ATAAGGAAACTGCTCTTTTACGCTGGTATGGACAGGTTAAGTTCAAGAGCAAGT 347

QY 339 AACCCAGGCGACAGATTAGAAATACCATTTTAGAGTCTTAAAGCATAAAGGCGCATGATCTG 398  
DB 348 TGTACCAAGGCGACCAATTTGTTATGACAGCGACTTTTGTAAAAACGTCGTGGCACCATAGC 407

QY 399 GCAAGTGGGTGGCAGCGCTCAAGTGGATGCGCAAGTGGTCTGC 440  
DB 408 TGTGGTTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 449

QY 219 CATGGCGCAAAACGGGAGGTTTATAGCTTCACATGCTTTGTTGGGGTTTGACCCCTGAAT 278  
DB 186 CTTTGGCGCAAACTGCCGCTGTGTGGAGTTATCAA-----AACCTGAAA- 229

QY 279 CGCCAAACCAAAATCGTGTATTTTCATGACGATTTAGATTAAGTTAAATTCGCGCATCCCTGT 338  
DB 230 --ATAAGGAAACTGCTCTTTTACGCTGGTATGGCAAGGTTAAGTTCAAGAGCAAGT 287

QY 339 AACCCAGGCGACAGATTAGAAATACCATTTTAGAGTCTTAAAGCATAAAGGCGCATGATCTG 398  
DB 288 TGTACCAAGGCGACCAATTTGGTTATGACAGCGACTTTTGTAAAAACGTCGTGGCACCATAGC 347

QY 399 GCAAGTGGGTGGCAGCGCTCAAGTGGATGCGCAAGTGGTCTGC 440  
DB 348 TGTGGTTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 389

RESULT 11  
US-10-089-019-17  
; Sequence 17, Application US/10089019  
; Patent No. 6951729  
; GENERAL INFORMATION:  
; APPLICANT: DEWOLF, WALTER E. JR  
; APPLICANT: KALLENDER, HOWARD  
; APPLICANT: LONSDALE, JOHN T.  
; TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID  
; FILE REFERENCE: GM50068  
; CURRENT APPLICATION NUMBER: US/10/089,019  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER: PCT/US00/29451  
; PRIOR FILING DATE: 2000-10-26  
; PRIOR APPLICATION NUMBER: 60/161,775  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 483  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-089-019-17

Query Match 15.8%; Score 75.6; DB 3; Length 483;  
Best Local Similarity 55.2%; Pred. No. 1.9e-14;  
Matches 222; Conservative 0; Mismatches 159; Indels 21; Gaps 3;

QY 39 TATAGAGCATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAT 98  
DB 69 TATTCAGGAATCAAGAGAGCTCTTCCCAACCGTTATCTCTATGCTTCTAGTGGACCGTGT 128

QY 99 TATAGAGTTACAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAAATGAAGA 158  
DB 129 CTTTGA---AGTGAGGAGGATACCATTTGCTATCAAAATGTGACCATCAACGAGCC 185

QY 159 CGTGTTTAACGGGCATTTCCCTAATAAGCCATTTTCCCGGCGTTTGTAGTCGTAGAGGG 218  
DB 186 TTTCTTTAACGGGCATTTCTCTCAATACCCAGTTATGCCAGTGTGTGTGATTATGGAAGC 245

QY 219 CATGGCGCAAAACGGGAGGTTTGTAGCTTCACTAGCTTTGGGGGTTTGACCCCTGAAT 278  
DB 246 CTTTGGCGCAAACTGCCGCTGTGTGGAGTTATCAA-----AACCTGAAA- 289

QY 279 CGCCAAACCAAAATCGTGTATTTTCATGACGATTTAGAGTCTTAAAGCATAAAGGCGCATGAT 338  
DB 290 --ATAAGGAAACTGCTCTTTTACGCTGGTATGGACAGGTTAAGTTCAAGAGCAAGT 347

QY 339 AACCCAGGCGACAGATTAGAAATACCATTTTAGAGTCTTAAAGCATAAAGGCGCATGATCTG 398  
DB 348 TGTACCAAGGCGACCAATTTGTTATGACAGCGACTTTTGTAAAAACGTCGTGGCACCATAGC 407

QY 399 GCAAGTGGGTGGCAGCGCTCAAGTGGATGCGCAAGTGGTCTGC 440  
DB 408 TGTGGTTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 449



US-09-902-540-3820

Query Match 15.3%; Score 73.4; DB 3; Length 462;  
Best Local Similarity 50.1%; Pred. No. 9.9e-14;  
Matches 219; Conservative 0; Mismatches 206; Indels 12; Gaps 1;  
QY 40 ATAGAGCATATCTTACAAATTCCTACCGCTATCCCATGCTTTTAGTGGATAGAATT 99  
DB 7 ATCGGAGAGATCTCAATCTGCTGCGCATCGGTACCGGTCTCTGCTGGTGGACCGGGTG 66  
QY 100 ATAGAGATTACAAGCCCAATAAAAAATTTGTGCTTTATAAGAAATATCACTTTTAAATGAAGAC 159  
DB 67 GTGGAGATCATCCCGGCCAGAGCTGACGGCTTACAGAACCTCACCATCAAGAGCCC 126  
QY 160 GTGTTTAAAGGGCATTTCCCTTAATAGCCCAATTTCCCGGGCGTTTGTATCGTAGAGGGC 219  
DB 127 TTCTTCAACGGGCACATTCGCGGTACCCCGGTGATCGCGGTGTGCTCATCTCTGGAGGCG 186  
QY 220 ATGGCGCAAAAGGGAGGGTTTATAGCTTCACTAGCTTTGTGGGGTTTGACCCCTGAAATC 279  
DB 187 CTGGCGCAGGCGAGCCCATCTCGCTTACAAGAGCGAACAATGGAT----- 234  
QY 280 GCCAAACAAAATCTGTATTTATGACGATTTGATAAGGTTAAATTCGCGATCCCTGTA 339  
DB 235 CGTCCGCGAAGCTCACGTACCTGATGGCGGTGGACGGCGCGCTTCCGCAAGCCGGTG 294  
QY 340 ACCCGAGCGCAGATFAGAAATACCAATTTAGAGTCTTAAAGCATAAAGGCGCATGATCTGG 399  
DB 295 TTGCGGGGGATCGCTCCAGTTGGAGATCGAGGTGGTGGTGCACAAAGGGCGCTGTCTGG 354  
QY 400 CAAAGTGGTGGCAGCGCTCAAGTGGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 459  
DB 355 AAGACGAAGGGCTGGCGCAGCTGCGACGGCGCGGTGTCGCGAAGGGCGAGTTCTCTGGCA 414  
QY 460 ATGATTGCAGAGAGAGA 476  
DB 415 ACCGTCTGGACAAGGA 431

RESULT 15  
US-09-902-540-1188  
; Sequence 1188, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIORITY FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIORITY FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1188  
; LENGTH: 21010  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(21010)  
; OTHER INFORMATION: unsure at all n locations  
US-09-902-540-1188

Query Match 15.3%; Score 73.4; DB 3; Length 21010;  
Best Local Similarity 50.1%; Pred. No. 7e-13;  
Matches 219; Conservative 0; Mismatches 206; Indels 12; Gaps 1;  
QY 40 ATAGAGCATATCTTACAAATTCCTACCGCTATCCCATGCTTTTAGTGGATAGAATT 99  
DB 12878 ATCGGAGAGATCTCAATCTGTCTGCGCATCGTACCGGTCTCTGCTGGTGGACCGGGTG 12937

QY 100 ATAGAGTTACAAGCCCAATAAAAAATTTGTGCTTATAGAAATATCACCTTTTAAATGAAGAC 159  
DB 12938 GTGGAGATCATCTCCCGGCCAGAGCTGACGGCTTACAGAAACGTCAACATCAACGAGCCC 12997  
QY 160 GTGTTTAAAGGGCATTTCCCTTAATAGCCCAATTTCCCGGGCGTTTGTATCGTAGAGGGC 219  
DB 12998 TTCTTCAACGGGCACATTCGCGGTACCCCGGTGATCGCGGTGTGCTCATCTCTGGAGGG 13057  
QY 220 ATGGCGCAAAAGGGAGGGTTTATAGCTTCACTAGCTTTGTGGGGTTTGACCCCTGAAATC 279  
DB 13058 CTGGCGCAGGCGACGGCCATCTCGCTTACAAGAGCGAGAACATGGAT----- 13105  
QY 280 GCCAAACAAAATCTGTATTTATGACGATTTGATAAGGTTAAATTCGCGATCCCTGTA 339  
DB 13106 CGTCCGCGAAGCTCACGTACCTGATGGCGGTGACCGCGCGCTTCCGCAAGCCGGTG 13165  
QY 340 ACCCGAGCGCAGATFAGAAATACCAATTTAGAGTCTTAAAGCATAAAGGCGCATGATCTGG 399  
DB 13166 TTGCGGGGGATCGCTCCAGTTGGAGATCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 13225  
QY 400 CAAAGTGGTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 459  
DB 13226 AAGACGAAGGGCTGGCGCAGCGTGCACGGCGCGGTGTCGCGAAGGGCGAGTTCTCTGGCA 13285  
QY 460 ATGATTGCAGAGAGAGA 476  
DB 13286 ACCGTCTGGACAAGGA 13302

Search completed: January 12, 2006, 03:14:46  
Job time : 156 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:21:48 ; Search time 658 Seconds  
(without alignments)  
6032.373 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atgggaacaaagcattcaaaaa.....tgattgcagagagatttaa 480

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA\_Main:

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 3: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 6: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 7: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 8: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 9: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 10: /cgn2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	480	100.0	480	7	US-10-282-122A-22798		Sequence 22798, A
2	480	100.0	480	9	US-10-662-126-36		Sequence 36, Appl
3	454.4	94.7	480	3	US-09-815-242-7504		Sequence 7504, Ap
4	454.4	94.7	480	7	US-10-335-977-2213		Sequence 2213, Ap
5	145.4	30.3	180	7	US-10-335-977-2212		Sequence 2212, Ap
6	134.4	28.0	441	7	US-10-282-122A-18085		Sequence 18085, A
7	99.2	20.7	459	6	US-10-369-493-35520		Sequence 35520, A
8	95	19.8	432	7	US-10-282-122A-16557		Sequence 16557, A
9	93.4	19.5	417	7	US-10-282-122A-21741		Sequence 21741, A
10	92.6	19.3	462	7	US-10-282-122A-41232		Sequence 41232, A
11	88.6	18.5	435	7	US-10-282-122A-24354		Sequence 24354, A
12	88.6	18.5	1755	7	US-10-398-221-1922		Sequence 1922, Ap
13	88.6	18.5	1927	7	US-10-398-221-3523		Sequence 3523, Ap
14	88.2	18.4	459	7	US-10-282-122A-31282		Sequence 31282, A
15	87.4	18.2	453	7	US-10-282-122A-14857		Sequence 14857, A
16	87	18.1	684707	7	US-10-398-221-9		Sequence 9, Appl
17	87	18.1	3011208	7	US-10-398-221-2058		Sequence 2058, Ap
18	86.4	18.0	426	3	US-09-815-242-6826		Sequence 6826, Ap
19	86.4	18.0	6021	3	US-09-070-927A-458		Sequence 458, App
20	86.2	18.0	414	6	US-10-369-493-35390		Sequence 35390, A
21	86.2	18.0	414	6	US-10-369-493-37892		Sequence 37892, A
22	86.2	18.0	414	6	US-10-369-493-38330		Sequence 38330, A
23	86.2	18.0	414	6	US-10-369-493-38792		Sequence 38792, A

24	86.2	18.0	426	7	US-10-282-122A-15658		Sequence 15658, A
25	85	17.7	423	7	US-10-282-122A-21199		Sequence 21199, A
26	85	17.7	471	6	US-10-369-493-23894		Sequence 23894, A
27	85	17.7	518	6	US-10-369-493-44892		Sequence 44892, A
28	83.6	17.4	420	6	US-10-369-493-44343		Sequence 44343, A
29	82.8	17.2	438	7	US-10-282-122A-33603		Sequence 33603, A
30	82.8	17.2	441	3	US-09-815-242-7839		Sequence 7839, Ap
31	82.8	17.2	441	7	US-10-282-122A-30332		Sequence 30332, A
32	82.8	17.2	441	9	US-10-965-006-1		Sequence 1, Appl
33	82.8	17.2	441	9	US-10-965-006-3		Sequence 3, Appl
34	82	17.1	435	6	US-10-369-493-33616		Sequence 33616, A
35	81.8	17.0	435	6	US-10-369-493-23694		Sequence 23694, A
36	81.6	17.0	441	6	US-10-369-493-41684		Sequence 41684, A
37	81.4	17.0	399	6	US-10-369-493-47014		Sequence 47014, A
38	80.2	16.7	518	7	US-10-282-122A-32389		Sequence 32389, A
39	78.8	16.4	423	7	US-10-282-122A-37625		Sequence 37625, A
40	78.8	16.4	432	7	US-10-282-122A-9293		Sequence 9293, Ap
41	78.8	16.4	3274	3	US-09-070-927A-350		Sequence 350, App
42	78.4	16.3	432	7	US-10-282-122A-6335		Sequence 6335, Ap
43	78.4	16.3	435	3	US-09-815-242-6420		Sequence 6420, Ap
44	78.2	16.3	453	7	US-10-282-122A-23419		Sequence 23419, A
45	78	16.2	435	7	US-10-282-122A-25130		Sequence 25130, A

ALIGNMENTS

RESULT 1

US-10-282-122A-22798  
; Sequence 22798, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA 034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22798  
; LENGTH: 480

```

; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-282-122A-22798

Query Match      100.0%; Score 480; DB 7; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.6e-131;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||
Db 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||

QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGAGCAATAAA 120
   |||
Db 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGAGCAATAAA 120
   |||

QY 121 AAAATTGTGCTTTATAGAATATCACTTTTAAATGAAGACGTTTAAACGGGCATTTCCCT 180
   |||
Db 121 AAAATTGTGCTTTATAGAATATCACTTTTAAATGAAGACGTTTAAACGGGCATTTCCCT 180
   |||

QY 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACAGGAGGGTTT 240
   |||
Db 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACAGGAGGGTTT 240
   |||

QY 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGAACCTCGAATCGCAAAACAGGAGGGTTT 300
   |||
Db 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGAACCTCGAATCGCAAAACAGGAGGGTTT 300
   |||

QY 301 TTATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGCAGACAGATTAGAA 360
   |||
Db 301 TTATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGCAGACAGATTAGAA 360
   |||

QY 361 TACCATTTAGAACTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420
   |||
Db 361 TACCATTTAGAACTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420
   |||

QY 421 GTGGATGGCAAGTGTGCTGTAAGCGCAATTGAAAGCCATGATTCAGAGAGAGATTAA 480
   |||
Db 421 GTGGATGGCAAGTGTGCTGTAAGCGCAATTGAAAGCCATGATTCAGAGAGAGATTAA 480
   |||

RESULT 3
US-09-815-242-7504
; Sequence 7504, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7504
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Helicobacter pylori

Query Match      100.0%; Score 480; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.6e-131;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||
Db 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||

QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGAGCAATAAA 120
   |||
Db 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGAGCAATAAA 120
   |||

QY 121 AAAATTGTGCTTTATAGAATATCACTTTTAAATGAAGACGTTTAAACGGGCATTTCCCT 180
   |||
Db 121 AAAATTGTGCTTTATAGAATATCACTTTTAAATGAAGACGTTTAAACGGGCATTTCCCT 180
   |||

QY 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACAGGAGGGTTT 240
   |||
Db 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACAGGAGGGTTT 240
   |||

QY 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGAACCTCGAATCGCAAAACAGGAGGGTTT 300
   |||
Db 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGAACCTCGAATCGCAAAACAGGAGGGTTT 300
   |||

QY 301 TTATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGCAGACAGATTAGAA 360
   |||
Db 301 TTATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGCAGACAGATTAGAA 360
   |||

QY 361 TACCATTTAGAACTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420
   |||
Db 361 TACCATTTAGAACTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420
   |||

QY 421 GTGGATGGCAAGTGTGCTGTAAGCGCAATTGAAAGCCATGATTCAGAGAGAGATTAA 480
   |||
Db 421 GTGGATGGCAAGTGTGCTGTAAGCGCAATTGAAAGCCATGATTCAGAGAGAGATTAA 480
   |||

RESULT 2
US-10-662-126-36
; Sequence 36, Application US/10662126
; Publication No. US20050063987A1
; GENERAL INFORMATION:
; APPLICANT: Knapp, Bernhard
; APPLICANT: Hundt, Brika
; APPLICANT: Schmidt, Karl-Heinz
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of
; FILE REFERENCE: CHIR-0340
; CURRENT APPLICATION NUMBER: US/10/662,126
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 09/230,158
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/IB97/00981
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(480)
; OTHER INFORMATION: bp mature protein; 17 kD protein from Helicobacter
; OTHER INFORMATION: pylori
US-10-662-126-36
```



```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(480)
US-09-815-242-7504

Query Match          94.7%; Score 454.4; DB 3; Length 480;
Best Local Similarity 96.7%; Pred. No. 1.3e-123;
Matches 464; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAAAT 60
DB 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAGAT 60
QY 61 CTACTCCAGCTATCCCATGCTTTTAGTGGATAGATTTATAGATTTACAGCCATATA 120
DB 61 TTACTCTACCGCTACCCCATGCTTTTAGTGGATAGATTTATAGATTTACAGCCATATA 120
QY 121 AAAATTTGCTTTATAGAAATATCACTTTTAAATGAAGACGTGTTTAAACGGGCAATTTCCCT 180
DB 121 AAAATTTGCTTTATAGAAATATCACTTTTAAATGAAGATGTTTAAACGGGCAATTTCCCT 180
QY 181 AATAAGCCCAATTTCCCGGCGGTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
DB 181 AATAAGCCCAATTTCCCGGCGGTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
QY 241 TTAGCCTTCACTAGCTTTGCGGGGTTTTGACCCCTGAAATCGCCAAAACAAAATTCGTGTAT 300
DB 241 TTAGCCTTCACTAGCTTTGCGGGGTTTTGACCCCTGAAATCGCCAAAACAAAATTCGTGTAT 300
QY 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGGACAGATTTAGAA 360
DB 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGGACAGATTTAGAA 360
QY 361 TACCAATTTAGAGTCTTAAAGCATAGGCGCATGCTCGCAAGTGGTGCGACCGCTCAA 420
DB 361 TACCAATTTAGAGTCTTAAAGCATAGGCGCATGCTCGCAAGTGGTGCGACCGCTCAA 420
QY 421 GTGATGCGCAAAAGTGTGCTGCTGAAGCGCAATTAAGAGCCATGATCGCAGAGAGATAA 480
DB 421 GTGATGCGCAAAAGTGTGCTGCTGAAGCGCAATTAAGAGCCATGATCGCAGAGAGATAA 480

RESULT 4
US-10-335-977-2213
; Sequence 2213, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
```

```
;
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...480
; SEQUENCE DESCRIPTION: SEQ ID NO: 2213:
US-10-335-977-2213

Query Match          94.7%; Score 454.4; DB 7; Length 480;
Best Local Similarity 96.7%; Pred. No. 1.3e-123;
Matches 464; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAAAT 60
DB 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAGAT 60
QY 61 CTACTCCAGCTATCCCATGCTTTTAGTGGATAGATTTATAGATTTACAGCCATATA 120
DB 61 TTACTCTACCGCTACCCCATGCTTTTAGTGGATAGATTTATAGATTTACAGCCATATA 120
QY 121 AAAATTTGCTTTATAGAAATATCACTTTTAAATGAAGACGTGTTTAAACGGGCAATTTCCCT 180
DB 121 AAAATTTGCTTTATAGAAATATCACTTTTAAATGAAGATGTTTAAACGGGCAATTTCCCT 180
QY 181 AATAAGCCCAATTTCCCGGCGGTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
DB 181 AATAAGCCCAATTTCCCGGCGGTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
QY 241 TTAGCCTTCACTAGCTTTGCGGGGTTTTGACCCCTGAAATCGCCAAAACAAAATTCGTGTAT 300
DB 241 TTAGCCTTCACTAGCTTTGCGGGGTTTTGACCCCTGAAATCGCCAAAACAAAATTCGTGTAT 300
QY 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGGACAGATTTAGAA 360
DB 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGGACAGATTTAGAA 360
QY 361 TACCAATTTAGAGTCTTAAAGCATAGGCGCATGCTCGCAAGTGGTGCGACCGCTCAA 420
DB 361 TACCAATTTAGAGTCTTAAAGCATAGGCGCATGCTCGCAAGTGGTGCGACCGCTCAA 420
QY 421 GTGATGCGCAAAAGTGTGCTGCTGAAGCGCAATTAAGAGCCATGATCGCAGAGAGATAA 480
DB 421 GTGATGCGCAAAAGTGTGCTGCTGAAGCGCAATTAAGAGCCATGATCGCAGAGAGATAA 480

RESULT 5
US-10-335-977-2212
; Sequence 2212, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
```

COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows NT 4.0  
SOFTWARE: UNIX  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/335,977  
FILING DATE: 30-Dec-2002  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/993,002  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragoras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: GTN-018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 2212:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 180 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...180  
SEQUENCE DESCRIPTION: SEQ ID NO: 2212:  
US-10-335-977-2212

Query Match 30.3%; Score 145.4; DB 7; Length 180;  
Best Local Similarity 92.2%; Pred. No. 2e-32;  
Matches 165; Conservative 0; Mismatches 11; Indels 3; Gaps 1;  
QY 161 TGTTTAACGGGCAATTCCTAATAAGCCATTTCCCGGCGGTTTGTGATGCTAGAGGGCA 220  
DB 2 TGTGTATGGGCAATTCCTAATAAGCCATTTCCCGGCGGTTTGTGATGCTAGAGGGCA 61  
QY 221 TGGCGCAACGGAGGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGACCTCGAATCG 280  
DB 62 TGGCGCAACGGAGGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGACCTCGAATAG 121  
QY 281 CCA--AAACAAAAATCGTGATTTTCATGACGATTGATAAGGTTAAATTCGCGATCCCT 336  
DB 122 CCAACAAAAAAATCCGTTGATTTTCATGACGATTGACAGGTTAAATTCGCGATTCCT 180

## RESULT 6

US-10-282-122A-18085  
Sequence 18085, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18085  
LENGTH: 441  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
US-10-282-122A-18085

Query Match 28.0%; Score 134.4; DB 7; Length 441;  
Best Local Similarity 58.4%; Pred. No. 5.7e-29;  
Matches 256; Conservative 0; Mismatches 176; Indels 6; Gaps 1;

QY 39 TATAGAGCATATCTTACAAATCTACCTCACCCTATCCCATGCTTTTACTAGTAGATAAT 98  
DB 9 TGTAAATGCAAAATTCAGAAATCTTACCACACCGTTATCCCTTTTACTAGTAGATAAT 68  
QY 99 TATAGAGTTCAAGCCCAATAAAAAATTTGCGCTTATAAGAAATATCACTTTTAATGAAGA 158  
DB 69 TACAGAAATTAAGTTTAAAGAGTGGTGCTAGGATATAAAAACATCAGTATAAGTGACCA 128  
QY 159 CGTGTTTTAAACGGGCAATTTCCCTAATAAGCCATTTTCCCGGCGGTTTGTATCGTAGAGG 218  
DB 129 TGTTTTATGGACATTTTCCAGGCGCATCTTATTTATCTCGAGTTTTGTATTTAGAAG 188  
QY 219 CATGGCGCAACGGAGGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGACCTCGAAT 278  
DB 189 TATGGCTCAAAACAGGTGGAGTTTGTAGCTTTTGAAGAT- - - - -GGAAGATAAAAGTGA 242  
QY 279 CGCCAAACAAATTCGTGTATTTTCATGACGATTGATAAGTTTAAATTCGCGATCCCTGT 338  
DB 243 TCCAAAAAAGTAAAGTAGTTTATTTTCAAGGATAGTAGAGCAAAATTTAGAAATCCCTGT 302  
QY 339 AACCCAGGCGACAGATTAGAAATACCATTTAGAAGTCTTTAAAGCATAAGGGCATGATCTG 398  
DB 303 CGCTCTGGAGATAGCTTGTATTAAGATGAGCGTGTAAATTCGTTAAATATGTG 362  
QY 399 GCAAGTGGGTGGCAGCGGCTCAAGTGGGATGGCAAGTGGTTCGCTGAAGCGCAATTGAAGC 458  
DB 363 GATTTTAAAGGGCAAGCTTTTGTAGATGGAATTTAGTTGACAGAGCGCGAGCTTTAAAGC 422  
QY 459 CATGATTCAGAGAGAGA 476  
DB 423 TATGATAGTGGATAAATA 440

## RESULT 7

US-10-369-493-35520  
Sequence 35520, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 35520  
; LENGTH: 459  
; TYPE: DNA  
; ORGANISM: Mesorhizobium loti  
US-10-369-493-35520

Query Match 20.7%; Score 99.2; DB 6; Length 459;  
Best Local Similarity 55.1%; Pred. No. 1.5e-18;  
Matches 225; Conservative 0; Mismatches 168; Indels 15; Gaps 1;

Qy	59	TTCTACCTCACCGCTATCCCATGCTTTTACTGGATAGTAATATAGAGTTCACAGCCCAATA	118
Db	53	TCCTGCACACCGCTATCCCTTCTGATGATCGACCGCATCATCGACATCGATGGTGAAG	112
Qy	119	AAAAAATTGCTGTTTAAAGATATACATTTTAAATGAAGACGTGTTTAAAGCGGCATTTCC	178
Db	113	ATTCCGCCATCGGCATCAAGACGTACCATCAAGCGCGCATTTTCAGGGCATTTCC	172
Qy	179	CTAATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGCGATGGCGCAACCGGGAGGT	238
Db	173	CGGAGCAGCGGCGTATGCGGGGCGTCTGATCGTGAAGCATGTCGCGCAGCGGAGCGG	232
Qy	239	TTTTAGCCTTCACTAGCTTGTGGGGGTTTGAACCTGMAATGCGCAAAACAAAATCGTGT	298
Db	233	CCATCTGCATCCGACGCTCGGGCGT-----CGAAGCCGTCGCTGCTCT	277
Qy	299	ATTTCATGACGATTGATAGTTTAAATTCGATCCCTGTAACCCGAGCGGACAGATTAG	358
Db	278	ATTTCGACCATCGACACGCGCAATTCGCAACCGGTCGTCGCGGCGACAGTTGA	337
Qy	359	AATACCAATTAGAGTCTTAAAGCATAGGCGCATGATCTGGCAAGTGGTGGCAGCGCTC	418
Db	338	AGATTCACTCAAAAAATCAAGACGCGGCAACCTGCTCAATTCGCTGTGAAGCCC	397
Qy	419	AAGTGATGGCAAAAGTGGTGGTGAAGCGCAATTTGAAGCCATGATG	466
Db	398	TGGTCGATGGCACCAGGCGGCGGCGGCGAGATTTTCAGCCATGATGG	445

## RESULT 8

US-10-282-122A-16557  
; Sequence 16557, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16557  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Clostridium botulinum  
US-10-282-122A-16557

Query Match 19.8%; Score 95; DB 7; Length 432;  
Best Local Similarity 53.5%; Pred. No. 2.5e-17;  
Matches 234; Conservative 0; Mismatches 185; Indels 18; Gaps 1;

Qy	34	TTTTTTATAGAGCATATCTTACAAATTTCTACTCACCCCTATCCCATGCTTTTAGTGAT	93
Db	13	TTAGATATAATGAATAAAAAAATTTATTCGCACAGATATCCATTTTATTAGTAGAT	72
Qy	94	AGATATTATAGAGTTTACAAGCCAATAAAAAAATTTGCTGCTTATAAGAAATATCATTTTAA	153
Db	73	AAAACTACTGAATTAGAAGAGGGAAGTGCAGTTGGATATAAAAAATGTACAGCTAAT	132
Qy	154	GAGACGTGTTTAAAGCGGCATTTCCCTAAATAAGCCCATTTTCCCGGGCGTTTGTATGTA	213
Db	133	GAATACCTTTTAAATGGGCATTTTCCGGAAGAACCAAGTAATGCTCGGGTTCGATTATA	192
Qy	214	GAGGCATGCGCAACCGGAGGGTTTGTAGCCCTTCACTAGCTTCTGGGGGTTTGACCT	273
Db	193	GAAGCTTTAGCACAGGTCGGAGCTGTTGCTATTTTAAAGTA-----AA	234
Qy	274	GAATCGCCAAACAAAAATCGTGTATTTTATGATGATTAAGGTTTAAATTCGCGCATC	333
Db	235	GAAGAGTTTAAAGGAATAATAGCTTATTTTGGAGGTATAAATAAGCTTAAGTTCAAAAA	294
Qy	334	CCTGTAAACCCAGCGCAGATTAGAAATACATTTAGAAGCTTTAAAGCATAAAGGCGATG	393
Db	295	AAGGTAGTACACAGAGATGTTTAAAACTTAGTAGTAGATCTTACTTAAATAAAGCGCT	354
Qy	394	ATCTGGCAAGTGGTGGCAGCGCTCAAGTGGATGGCAAGTGGTGGTGAAGCCGAATTG	453
Db	355	GCAGGAGTAGGTAAGCGCGTAGCTACTGTAGATCGGAAAGTAGTGTGCAAGACAGATTA	414
Qy	454	AAAGCCATGATTGCAGA	470
Db	415	TTATTTGTAATAGGAAA	431

## RESULT 9

US-10-282-122A-21741  
; Sequence 21741, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos

```

; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21741
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Enterococcus faecium
; US-10-282-122A-21741

Query Match          19.5%; Score 93.4; DB 7; Length 417;
Best Local Similarity 53.6%; Pred. No. 7.4e-17;
Matches 230; Conservative 0; Mismatches 181; Indels 18; Gaps 1;

QY 40 ATAGAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAATT 99
DB 7 ATACAAGAAATAAAGAAATCATTTCCACATCGTTATCCAATGCTTTTGATTGATCGAGTG 66

QY 100 ATAGAGTTACAGCCCAATAAAAAATTTGCGCTTATAAGAAATACACTTTTAATGAAGAC 159
DB 67 GAAGAAATGGTTGAAGGAGAACGGATCGTTGCTTAAGAAAAATGTAACGATCAATGAGCCC 126

QY 160 GGTGTTTAACGGCATTTCCCTATATAGCCCATTTTCCCGGGCGTTTGTATCGTAGAGGCG 219
DB 127 TTTTTCAGGAGCATTTTCCAGAAACCTGTTATGCCAGGGGTATGATTGTTGAAGCA 186

QY 220 ATGGCGCAACACGGGAGGGTTTTTAGCCTTCACATGCTTGTGGGGGTTTGACCCCTGAAATC 279
DB 187 ATGGCTCAAGCAGGAGCAGTTGCACTATTCTTCT- - - - -CTTGAACAA 228

QY 280 GCCAAACAAAAATCGTGATTTTTCATGACGATTTGAAGTTAAATTCGCGATCCCTGTGA 339
DB 229 TTCAAGGGAAAGACGGCTTATTTCCGCGGATTCGACAAAGCAAAATTCGTAAGAAAGTT 288

QY 340 ACCCGAGGACAGATAGNATACCATTTAGAGTCTTAAGCATTAAGGCGCATGATCTGG 399
DB 289 ACACCGGGAGATACCTTATACCTTGAAGTTGAGATATTGAAGTGAAGGCTTCAGCGCGGA 348

QY 400 CAAAGTGGTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTCTGAAAGCGCAATTTGAAAGCC 459
DB |||||
```

```

Db 349 ATAGGCAAGGAATCGCTAAAGTAGACGGTAAAAAAGTAGCCGAGCGGAATTACCTTT 408
QY 460 ATGATTGCA 468
Db 409 ATGATTGGA 417

RESULT 10
US-10-282-122A-41232
; Sequence 41232, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41232
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Vibrio cholerae
; US-10-282-122A-41232

Query Match          19.3%; Score 92.6; DB 7; Length 462;
Best Local Similarity 53.8%; Pred. No. 1.4e-16;
Matches 222; Conservative 0; Mismatches 179; Indels 12; Gaps 1;

QY 59 TTCTACTCTACCGCTATCCCATGCTTTTAGTGGATAGAATTTATAGAGTTACAAGCCAATA 118
DB 56 TTCTGCTCTACCGCTATCCCATTTTACTGATGATCGGTCGCCGACTATGAAGAGTA 115

QY 119 AAAAAAATGTGCTTTAAGAATATCACTTTTAATGAAGACGTGTTTAACCGGGCAATTTCC 178
DB 116 AGTACCTGATGTTGTTGAAAAATGTTTCAGTGAATGAGCCTCAGTTTACGGGTCACTTTC 175

QY 179 CTAAATAGCCCAATTTTCCCGGGCGGTTTTTGATCGTAGAGGCGCATGGCGAAACGGGAGGT 238
DB 176 CACAACCTGCCGGTTTTTCCCGGCTGATTGATTTTAGAAGCCATGGCTCAAGCCACTGTGTC 235
```

Qy	239	TTTTAGCCTTCAC	TAGCTTGTGGGGTTGAC	CCCTGAAATCGCCAAACAA	AAAAATCGTGT	298
Db	236	TGTTGGGCTTCA	AAACCTTTCGTTGCGCCAA	-----AGAGAATGA	ACTGTACT	283
Qy	299	ATTTTCATGACGA	TATGATAAGGTTAAATTC	CCGCATCCCTGTAACCC	CAGCGACAGATTAG	358
Db	284	ATTTTGCAGCA	TATGATGAGCTTAAGTTC	CGTAAGCCAGTGAAC	CGGCGGATCAACTGA	343
Qy	359	AATACCA	TTTGAAGAGTCTTA	AAGCATAAAGGGCATGAT	CTTGCGCAAGTGGGTGGC	CAGGCTC 418
Db	344	TGGTTGAAGTT	GAAATTCCTTAAGAGCGT	CGTGGTATCGCGTGT	TTAACGCGGTGCCA	403
Qy	419	AAGTGGATGGC	AAAGTGTGCTCGCTG	TAAGCCGAATTGAA	AAGCCATGATTCAGAG	471
Db	404	AAGTFCGATGG	TGATGTGTCTGTTCT	GTCTCAACTGAAAT	TGCCAGCTCGAGAG	456

```

RESULT 11
US-10-282-122A-24354
; Sequence 24354, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```

Query Match 18.5%; Score 88.6; DB 7; Length 435;  
Best Local Similarity 64.3%; Pred. No. 2e-15;  
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 34 TTTTATAGAGCATATCTTCAAAATCTCTACCTACCGCTATCCCATGCTTTTATGATGAT 93

```

Db      4   TTAGATATTAAAGAAATCAAGAGAGATTTTGGCTCATCTGTTATCCATTTTGTGTAGTAGAT   63
Qy      94   AGAATTATAGAGTTTACAGGCCAATAAAAAAATGTGCGCTTATAGAATATCACITTTTAAAT   153
Db     104   TTTTCTGTTGGAAGGCGAAAAAGTTACCGCTATTAGAATGTAACAGCCCAAT   123
Qy     154   GAAGACGTGTTTAACGGGCAATTCCTTAATAAGGCCAATTTTCCCGGGCGTTTGTGATCGTA   213
Db     124   GAAGAAATTTTTTAACGGGCACTTTTCCCTGAGTATCCTGTAATGCCAGGAGTATTATAAGTG   183
Qy     214   GAGGCGATGGCGCAACGAGGAGGGTTT   240
Db     184   GAAGCTCTGGCGCAACAAAGTGAATTT   210

RESULT 12
US-10-398-221-1922
; Sequence 1922, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1922
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Listeria monocytogenes-4B
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide a,g,c or t/u
US-10-398-221-1922

```

Query Match	18.5%;	Score 88.6;	DB 7;	Length 1755;
Best Local Similarity	64.3%;	Pred. No. 4.2e-15;		
Matches 133;	Conservative 0;	Mismatches 74;	Indels 0;	Gaps 0;
Qy	34	TTTTTTATAGAGCATATCTTACAAAATCTACCTCACCGCTATCCCATGCTTTTAAGTGAT	93	
Db	830	TTAGATATTAAAGAAATCAAGAGATTTTGCCTCATCGTTATCCATTTTCTTAGTAGAT	889	
Qy	94	AGAATTATAGAGTTTACAAGCAATAAAAAATTTGTGCTTTAAGAATATCATCTTTTAAAT	153	
Db	890	AGAGTTATTCTGTTGAAGAGGCAAAAAAGTTACCGCTATTAAAGATGTAAACAGCAAAAT	949	
Qy	154	GAAGACGTGTTTAAACGGGCATTTTCCCTTAATAAGCCCATTTTCCGGCGCTTTTTCATCGTA	213	
Db	950	GAAGAAATTTTAAACGGGCATTTTCTTGAGTATCTCTGTAATGCGCAGAGTTTAAATAGTG	1009	
Qy	214	GAGGGCATCGCCAAACGGGAGGTTT	240	
Db	1010	GAAGCTCTAGCGCAAACTAGTGGAAAT	1036	

```

RESULT 13
US-10-398-221-3523
; Sequence 3523, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398.221

```

; CURRENT FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3523  
; LENGTH: 1927  
; TYPE: DNA  
; ORGANISM: Listeria monocytogenes 4b  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(end)  
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u  
US-10-398-221-3523

Query Match 18.5%; Score 88.6; DB 7; Length 1927;  
Best Local Similarity 64.3%; Pred. No. 4.4e-15;  
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
Qy 34 TTTTATAGACATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGAT 93  
Db 990 TTAGATATTAGAAATCAAGAGATTTTGCCTCATCGTTATCCATTTTGTAGTAGAT 1049  
Qy 94 AGAATTATAGACTTACAAGCCCAATAAAAAATTTGCTGTATAGAAATATCACTTTTAAT 153  
Db 1050 AGAGTTATTCTGTTGAAGAGCAAAAGTTACCGCTATTAAAGATGTAAACGCAAT 1109  
Qy 154 GAAGAGCTTTTAAAGCGCATTTCCCTAATAAGCCCATTTTCCCGGCGCTTTTATCGTA 213  
Db 1110 GAAGAAATTTTAAAGCGGCACTTCTGAGTATCTGTAAATGCCAGGAGTTTAAATAGTG 1169  
Qy 214 GAGGCGATGGCGCAACGGGAGGTTT 240  
Db 1170 GAAGCTCTAGCGCAAACTAGTGAATT 1196

## RESULT 14

US-10-282-122A-31282  
; Sequence 31282, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31282  
; LENGTH: 459  
; TYPE: DNA  
; ORGANISM: Pasteurella multocida  
US-10-282-122A-31282

Query Match 18.4%; Score 88.2; DB 7; Length 459;  
Best Local Similarity 53.3%; Pred. No. 2.7e-15;  
Matches 220; Conservative 0; Mismatches 178; Indels 15; Gaps 1;  
Qy 42 AGAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGATAT 101  
Db 39 AAATGAAATTTATGAAATTTGCTTCTCACCGTTATCTCTTTTATTAGTGGATCGTGTGT 98  
Qy 102 AGATTACAGCAATAAAAAATTTGCGCTTATAGAAATATCACTTTTAATGAAGAGT 161  
Db 99 GGATTATGAAGAGGTAATGGCTTAAAGCTGTAAAAAATATTAGCGTAAATAGCCCTTG 158  
Qy 162 GTTTAAGGGCATTTCCCTAATAAGCCCATTTTCCCGGCGCTTTTGATCGTAGAGGCGAT 221  
Db 159 CTTTACGGGACATTTCCAGACAACCGATTTTCCGGGTGTGTGATTTTGGAGCCAT 218  
Qy 222 GCGCAACCGGGAGGTTTTCCTTACCTTCTAGCTTGTGGGGTTTGACCTGAAATCGC 281  
Db 219 GGCACAAGCGAGCGGGTGTCTGCTGTAAAAAATTACCGTCAATTAG----- 265  
Qy 282 CAAAAAATAATCGTGTATTTTATGAGATTTAGAGTTAAATTCGATCCCTGTAAC 341  
Db 266 --AAATGAAATTTTATTTTGTGCAATTTGATGCGCGCTTAAACGTCCTGTGT 323  
Qy 342 CCAGGCGACAGATTAGAAATACCATTTTAGAAGTCTTAAAGCATAAGGCGCATGTGCA 401  
Db 324 GCCTGGTGATCAAAATGGTGTAGAGTTCACTTTTGAAGAGCGTCCGGGTATTACGG 383  
Qy 402 AGTGGGTGGCACCGCTCAAGTGGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGT 454  
Db 384 TTTCACTGGTGGCAACCGTAGATGCTCAAGTGGTGGTGGTGGTGGTGGTGGTGGT 436

## RESULT 15

US-10-282-122A-14857  
; Sequence 14857, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14857
; TYPE: DNA
; LENGTH: 453
; ORGANISM: Bordetella pertussis
US-10-282-122A-14857

Query Match      18.2%; Score 87.4; DB 7; Length 453;
Best Local Similarity 50.4%; Pred. No. 4.6e-15;
Matches 214; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

QY 40 ATAGAGCATATCTTACAAATTTACCTACCGCTATCCCATGCTTTTGTGGATAGAAAT 99
Db 13 ATCAAGGGATCATGGACCGGCTGCCGATCGCTACCCGATGCTGTCGACCGTG 72

QY 100 ATAGAGTTACAAGCCAATAAATAATTTGCTTTATAAGAAATATCACTTTTAAATGAAGAC 159
Db 73 CTGGAGATGGTCCCGGCAATCCATCGTTGCCATCAAGAACGCTCTCGATCAACGAGCCG 132

QY 160 GTGTTTAAACGGGCATTTCCTAATAAGCCCATTTTCCCGGGCGTTTGTGATCGTAGAGGGC 219
Db 133 TTTTTCACGGGCCATTTTCGGCATCATCCGGTCATGCGGGCGTGTGATCGTCGAGGCC 192

QY 220 ATGGCGCAAAACGGGAGGGTTTTTAGCCTTCACTAGCTTGTGGGGGTTTGACCCCTGAATC 279
Db 193 ATGGCGCAGGCTTCGGCCCTGTTCTCGTTCAACGACGAGAAACGGCGGGCTGAAGTGGAC 252

QY 280 GCCAAACAAAATCGTGTATTTTCATGACGATTGATAAGTTAAATTCGCATCCCTGTA 339
Db 253 GCGGCCAAGACCGCCTACTACTGTGTGGCATCGACGGCGCGGTTCCGTAAACCCGTG 312

QY 340 ACCCAGGCGACAGATTAGAATACCAATTTAGAAAGTCTTAAAGCATAAAGGCATGATCTGG 399
Db 313 GTGCCGGGCGACCAAGTTGGCGGTGGAGTCGAGGCCGAGGCCCTGTGCGTACCATCTGC 372

QY 400 CAAGTGGGTGGCAACCGCTCAAGTGGATGGCAAGTGGTCTGTAAGCCGGAATTTGAAAGCC 459
Db 373 AAGTACCGAGGCGGCGCCCTGTGTGACGCGCAACTGTTGCGCGAAGCCAAAGCTGATGTGC 432

QY 460 ATGAT 464
Db 433 GCGAT 437
```

Search completed: January 12, 2006, 03:25:50  
Job time : 659 secs



November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New). Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:29:32 ; Search time 226 Seconds  
(without alignments)  
1718.970 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggaacaaagccatcaaaa.....tgattgcagagagagattaa 480

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.New.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq2.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	73.6	15.3	447	6	US-10-467-657-6109 Sequence 6109, Ap
2	73.6	15.3	447	6	US-10-467-657-7549 Sequence 7549, Ap
3	73.6	15.3	552	6	US-10-467-657-6129 Sequence 6129, Ap
c 4	73.6	15.3	552	6	US-10-467-657-7559 Sequence 7559, Ap
c 5	60.8	12.7	2926	6	US-10-793-626-3747 Sequence 3747, Ap
6	35.8	7.5	1458	6	US-10-750-185-26969 Sequence 26969, A
7	35.8	7.5	1458	6	US-10-750-623-26969 Sequence 26969, A
8	35.6	7.4	822	6	US-10-793-626-727 Sequence 727, App
9	35.6	7.4	3215	6	US-10-793-626-3566 Sequence 3566, Ap
c 10	35.6	7.4	3993	6	US-10-793-626-3985 Sequence 3985, A
c 11	34.2	7.1	2528	6	US-10-750-185-51435 Sequence 51435, A
c 12	34.2	7.1	2528	6	US-10-750-623-51435 Sequence 51435, A
c 13	34	7.1	134499	7	US-11-117-187-192 Sequence 192, App
14	33.6	7.0	2382	6	US-10-750-185-57788 Sequence 57788, A
15	33.6	7.0	2382	6	US-10-750-623-57788 Sequence 57788, A
c 16	33.6	7.0	394468	7	US-10-995-561-13473 Sequence 13473, A
17	33.4	7.0	856	6	US-10-750-185-60462 Sequence 60462, A
18	33.4	7.0	856	6	US-10-750-623-60462 Sequence 60462, A
c 19	33.4	7.0	1304	6	US-10-750-185-52022 Sequence 52022, A
c 20	33.4	7.0	1304	6	US-10-750-623-52022 Sequence 52022, A
c 21	33.4	7.0	6182	6	US-10-240-708-87 Sequence 87, Appl
c 22	33	6.9	555	7	US-11-128-061-7215 Sequence 7215, Ap
c 23	33	6.9	2372	7	US-11-128-061-3573 Sequence 3573, Ap

ALIGNMENTS

RESULT 1

US-10-467-657-6109 7 600 6.8 32.8 6.8 600 7 US-11-136-527-5665 Sequence 5665, Ap  
; Sequence 6109, Application US/10467657 Sequence 39282, A  
; Publication No. US20050260581AI Sequence 39282, A  
; GENERAL INFORMATION: Sequence 1569, Ap  
; APPLICANT: CHIRON SpA Sequence 53522, A  
; APPLICANT: FONTANA Maria Rita Sequence 53522, A  
; APPLICANT: PIZZA Mariagrazia Sequence 396, App  
; APPLICANT: MASIGNANI Vega Sequence 59973, A  
; APPLICANT: MONACI Elisabetta Sequence 59973, A  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS Sequence 51969, A  
; FILE REFERENCE: Sequence 55313, A  
; CURRENT APPLICATION NUMBER: US/10/467,657 Sequence 55313, A  
; CURRENT FILING DATE: 2003-08-11 Sequence 13264, A  
; PRIOR APPLICATION NUMBER: GB-0103424.8 Sequence 31, Appl  
; PRIOR FILING DATE: 2001-02-12 Sequence 33663, A  
; NUMBER OF SEQ ID NOS: 9218 Sequence 47750, A  
; SOFTWARE: SeqWin99, version 1.04 Sequence 47750, A  
; SEQ ID NO 6109 Sequence 3384, Ap  
; LENGTH: 447 Sequence 32382, A  
; TYPE: DNA Sequence 13293, A  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-6109

Query Match	15.3%	Score 73.6	DB 6	Length 447
Best Local Similarity	51.2%	Pred. No. 1.7e-13		
Matches	206	Conservative	0	Mismatches 184; Indels 12; Gaps 1
Qy	47	ATATCTTACAAATCTACCTCACCGCTATCCATCGCTTTTAGTGGATAGAATTATAGAT	106	
Db	32	ACATCCAAAACCTCATCCCGACCGCTACCGCTTCTCCAGCTCGACCGCATACCGCT	91	
Qy	107	TACAAGCCAAATAAAAAATTTGCTTTATATAAGATATACATTTTAAATGAAGCGTGTFTA	166	
Db	92	TCGAGCGCATGMAAACCTTGACCGCCATCAAAAAGTAACCAACGAACCCCAATTCC	151	
Qy	167	ACGGGATTTCCCTTAATAAGCCCATTTTCCCGGGGTTTTCATCGTAGAGCGCATGGCG	226	
Db	152	AAGGCCATTTCCCGACCTGCGCGTTATGCGCGGTACTCATCATCGAAGCGATGGCG	211	
Qy	227	AAAGGGAGGAGTTTATAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCGCAAAA	286	
Db	212	AGCGTGGCGACGTTGGCGATTTTGAGCGAAGCGGGG-----CGCAAGGAAA	259	
Qy	287	CAAAAATCGTGTATTTTCATGACGATTTAATTCGGCATCCCTGTAAACCCAG	346	

Db 260 ACGAATTTTCTTCTTCGCGCGGATAGACGAAGCCGCTTTTCAACGCCCAAGTCATCCCCG 319  
Qy 347 GCGACAGATTAGAAATACCAATTTAGAAAGTCTTAAAGCATAAAGGCGATGATCTGGCAAGTGG 406  
Db 320 GCGACCAACTCGTCTTTGAGTCGAATCTCTGACNAGCGGCGGCGATCGGCAAAATTC 379  
Qy 407 GTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTGCCTGAAGCCG 448  
Db 380 ACGCCGTTGCCAAAGTGGACGACGAAGTGCCTGGAAGCCG 421

## RESULT 2

US-10-467-657-7549  
; Sequence 7549, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 7549  
; LENGTH: 447  
; TYPE: DNA  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-7549

Query Match 15.3%; Score 73.6; DB 6; Length 447;  
Best Local Similarity 51.2%; Pred. No. 1.7e-13;  
Matches 206; Conservative 0; Mismatches 184; Indels 12; Gaps 1;  
Qy 47 ATATCTTACAAATTTACCTCACCGCTATCCCATGCTTTTAGTGATAGAAATATAGAGT 106  
Db 32 ACATCAAAAACCTCATCCCCACCGCTACCGCTTTCTCCAGCTCGACCGCATTACCGCCT 91  
Qy 107 TACAAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAATGAAGAGTGTTTA 166  
Db 92 TCGAGCCGATGAACCCCTGACCGCATCAAAAACGTAAACCAATAAACGCCAAATTC 151  
Qy 167 ACGGCGAATTTCCCTAATAAGCCCAATTTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGC 226  
Db 152 AAGGCCAATTTCCCGACCTGCGCGTATGCGCGGCTACTCATCATCGAAGCGATGCGC 211  
Qy 227 AAACGGAGGGTTTTAGCTTCACTAGCTTTGTGGGGGTTTGAACCTCGTGAAGTGG 286  
Db 212 AGCGGTGCGGACGTTGGCGATTTTGAGCGAAGCGCGG-----CGCAAGGAAA 259  
Qy 287 CAAAATCGTGATTTTCATGACCAATGATAGTAAATTCGCCATCCCTGTAAACCCAG 346  
Db 260 ACGAATTTTCTTCTTCGCGGATAGACGACCGCTTTCAACAGCGCAAGTCATCCCCG 319  
Qy 347 GCGACAGATTAGAAATACCAATTTAGAAAGTCTTAAAGCATAAAGGCGATGATCTGGCAAGTGG 406  
Db 320 GCGACCAACTCGTCTTTGAAGTCGAATCTCTGACCGCGCGCGGCGATCGGCAAAATTC 379  
Qy 407 GTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTGCCTGAAGCCG 448  
Db 380 ACGCCGTTGCCAAAGTGGACGACGAAGTGCCTGCGCGCTGGAAGCCG 421

## RESULT 3

US-10-467-657-6129  
; Sequence 6129, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:

; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 6129  
; LENGTH: 552  
; TYPE: DNA  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-6129

Query Match 15.3%; Score 73.6; DB 6; Length 552;  
Best Local Similarity 51.2%; Pred. No. 2e-13;  
Matches 206; Conservative 0; Mismatches 184; Indels 12; Gaps 1;  
Qy 47 ATATCTTACAAATTTACCTCACCGCTATCCCATGCTTTTAGTGATAGAAATATAGAGT 106  
Db 47 ACATCAAAAACCTCATCCCCACCGCTACCGCTTTCTCCAGCTCGACCGCATTACCGCCT 106  
Qy 107 TACAAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAATGAAGAGTGTTTA 166  
Db 107 TCGAGCCGATGAACCCCTGACCGCATCAAAAACGTAAACCAATAAACGCCAAATTC 166  
Qy 167 ACGGCGAATTTCCCTAATAAGCCCAATTTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGC 226  
Db 167 AAGGCCAATTTCCCGACCTGCGCGTATGCGCGGCTACTCATCATCGAAGCGATGCGC 226  
Qy 227 AAACGGAGGGTTTTAGCTTCACTAGCTTTGTGGGGGTTTGAACCTCGTGAAGTGGCAAAA 286  
Db 227 AGCGGTGCGGACGTTGGCGATTTTGAGCGAAGCGCGG-----CGCAAGGAAA 274  
Qy 287 CAAAATCGTGATTTTCATGACCAATGATAGTAAATTCGCCATCCCTGTAAACCCAG 346  
Db 275 ACGAATTTTCTTCTTCGCGGATAGACGACCGCTTTCAACAGCGCAAGTCATCCCCG 334  
Qy 347 GCGACAGATTAGAAATACCAATTTAGAAAGTCTTAAAGCATAAAGGCGATGATCTGGCAAGTGG 406  
Db 335 GCGACCAACTCGTCTTTGAAGTCGAATCTCTGACCGCGCGCGGCGATCGGCAAAATTC 394  
Qy 407 GTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTGCCTGAAGCCG 448  
Db 395 ACGCCGTTGCCAAAGTGGACGACGAAGTGCCTGCGCGCTGGAAGCCG 436

## RESULT 4

US-10-467-657-7559/c  
; Sequence 7559, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 7559  
; LENGTH: 552  
; TYPE: DNA

; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-7559

Query Match 15.3%; Score 73.6; DB 6; Length 552;  
Best Local Similarity 51.2%; Pred. No. 2e-13;  
Matches 206; Conservative 0; Mismatches 184; Indels 12; Gaps 1;

QY 47 ATATCTTACAAATCTACCTACCGCTATCCCTGCTTTTAGTGGATAGATATAGAGT 106  
DB 506 ACATCCAAAACTCATCCCCACCGCTACCGCTTTCTCCAGCTCGACCGCATACCGCCT 447  
QY 107 TACAAGCCAAATAAAATTTGTCCTTATAGATATACCTTTTAATGAAGAGTGTTTA 166  
DB 446 TCAGCGGATGAACCTGACCGCATCAAAACGTAAACATACAGCAACCCCAATTC 387  
QY 167 ACGGCAATTCCTTAATAAGCCATTTTCCCGCGCTTTTGAATGATAGAGGCGATGCGC 226  
DB 386 AAGGCCAATTTCCCGCACTCCCGTATATGCGCGGTACTCATCGAAGCGATGCGC 327  
QY 227 AAACGGGAGGTTTTAGCTTCACTAGCTTGTGGGGTTTGACCTGAATCGCCAAA 286  
DB 326 AGCGGTGCGGCACGTTGGCGAATTTTGAGCGAAGCGGG-----CGCAAGGAAA 279  
QY 287 CAAATATCGTATTTTCATGACGATTGATAAGGTAAATTCGCACTCCCTGTAAACCCAG 346  
DB 278 ACGAATTTTCTTCTTCCCGCGCATAGACGACCGCTTTCAACGCGCAAGTATCCCG 219  
QY 347 GCGACAGATTAGAAATACCAATTTAGAAGTCTTAAAGCATAGGCGCATGATCTGCGAAGTGG 406  
DB 218 GCGACCAACTCGTCTTTGAATCGAATCTCTGACCGCGCGGCGCATCGGCAATTTCA 159  
QY 407 GTGGCAGCGCTCAAGTGGATGGCAAGTGTGCTGCTGAAGCG 448  
DB 158 ACGCCGTGTGCAAGTGGACGAGCAAGTGTGCGCGTGTGCAAGCG 117

RESULT 5

US-10-793-626-3747/c  
; Sequence 3747, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3747  
; LENGTH: 2926  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3747

Query Match 12.7%; Score 60.8; DB 6; Length 2926;  
Best Local Similarity 56.5%; Pred. No. 6.1e-09;  
Matches 113; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 18 AACCTTGCATCTCAATTTTATAGCATATCTTACAAATCTACCTACCGCTATCC 77  
DB 200 AAAATGGAAACGATTTTGAATTAATCAATTAACAAATTTATCCCATAGAACCC 141  
QY 78 CATGCTTTTAGTGGATAGATATAGATTACAGCAATAAAATTTGCTTATAA 137  
DB 140 TTTTCTACTAATAGATAAATTTGATAGATAGAGGATAAAGTGTGTAGGACTAA 81  
QY 138 GAATATCACTTTTAATGAAGAGCGTTTAAACGGGCAATTTCCCTTAATAAGCCATTTCC 197

DB 80 GCAAGTATCAGGTAATGAACCAATTTTCCAGGACATTTCCCTAATTAAGCTGTGATGCC 21  
QY 198 GGGCGTTTGTATCGTAGAGG 217  
DB 20 TGGTGTCTCTTATTATACAGAAG 1

RESULT 6

US-10-750-185-26969  
; Sequence 26969, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 26969  
; LENGTH: 1458  
; TYPE: DNA  
; ORGANISM: Bovine 19866881740741  
US-10-750-185-26969

Query Match 7.5%; Score 35.8; DB 6; Length 1458;  
Best Local Similarity 55.1%; Pred. No. 0.36;  
Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 25 CAATCTCAATTTTATAGACATATCTTACAAATTTCTACCTACCGCTATCCCAATGCTT 84  
DB 572 CTAAGTCCTCTGTTTAAATTAAGTGTACGATGTCACAAAGATTTTCCATATCTCTTGATGGCTT 631  
QY 85 TTAGTGGATAGATATAGATTACAGCCCAATTAATAAATTTGCTTATAAGATATC 144  
DB 632 CTAATGAATAAATTTTGTAGTGAAGCCCTATAAATAAATGTTTAAATAAATTTGCC 691  
QY 145 ACTTTTA 151  
DB 692 ATTITAA 698

RESULT 7

US-10-750-623-26969  
; Sequence 26969, Application US/10750623  
; Publication No. US20050287531A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-1  
; CURRENT APPLICATION NUMBER: US/10/750,623  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 26969  
; LENGTH: 1458  
; TYPE: DNA

```
; ORGANISM: Bovine 19866881740741
US-10-750-623-26969

Query Match          7.5%; Score 35.8; DB 6; Length 1458;
Best Local Similarity 55.1%; Pred. No. 0.36;
Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 25 CAATCTCAATTTTATAGAGCATATCTTACAAATCTACCTACCGCTATCCCATGCTT 84
   |||||
Db 572 CTAAGTCCCTGTTTAAATACGTGTCACAAAGATTTCCATATCTCTTGATTGGCTT 631
   |||||

QY 85 TTAGTGGATAGATTATAGATTACAGCCATAAAAAATTTGCTTTATAGATATC 144
   |||||
Db 632 CTAATGAATAAAATTTTGGTGAAGAGCCCTATAAAATAAATGTTTAATAATAATTTGCC 691
   |||||

QY 145 ACTTTTA 151
   |||||
Db 692 ATTTTAA 698

RESULT 8
US-10-793-626-727
; Sequence 727, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 727
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-727

Query Match          7.4%; Score 35.6; DB 6; Length 822;
Best Local Similarity 50.0%; Pred. No. 0.3;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 8 AAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATCTACCTC 67
   |||||
Db 449 AAAGTGTGTAATAATAAACTGCATAATTTAAGACAGAAAAATATTGAAATTTGACTG 508
   |||||

QY 68 ACCGCTATCCCATGCTTTTAGTGGATAGATTTATAGATTACAGCCATAAAAAATTTG 127
   |||||
Db 509 AAGATGTTGCAACGAGTTTCCAAAATAGTGTGTAGAAAGTTTAAACCTATAAAGCTATTC 568
   |||||

QY 128 TCGCTTATAGATATACATCTTTTAATGAACGCGTGTTTAACGGGCATTTCCCTAATAA 185
   |||||
Db 569 ATGCTTGTAAACTTATAATGTTAATCGCTTAATGTTGCAGGTGGTGTGCTAGTAA 626
   |||||

RESULT 9
US-10-793-626-3566
; Sequence 3566, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3566
; LENGTH: 3215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3566

Query Match          7.4%; Score 35.6; DB 6; Length 3215;
Best Local Similarity 50.0%; Pred. No. 0.66;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 8 AAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATCTACCTC 67
   |||||
Db 110 AAAGTGTGTAATAATAAACTGCATAATTTAAGACAGAAAAATATTGAAATTTGACTG 169
   |||||

QY 68 ACCGCTATCCCATGCTTTTAGTGGATAGATTTATAGATTACAGCCATAAAAAATTTG 127
   |||||
Db 170 AAGATGTTGCACGAGTTTCCAAAATAGTGTGTAGAAAGTTTAAACCTATAAAGCTATTC 229
   |||||

QY 128 TCGCTTATAGATATACATCTTTTAATGAACGCGTGTTTAACGGGCATTTCCCTAATAA 185
   |||||
Db 230 ATGCTTGTAAACTTATAATGTTAATCGCTTAATGTTGCAGGTGGTGTGCTAGTAA 287
   |||||

RESULT 10
US-10-793-626-3985/c
; Sequence 3985, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3985
; LENGTH: 3993
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3985

Query Match          7.4%; Score 35.6; DB 6; Length 3993;
Best Local Similarity 50.0%; Pred. No. 0.75;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 8 AAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATCTACCTC 67
   |||||
Db 1449 AAAGTGTGTAATAATAAACTGCATAATTTAAGACAGAAAAATATTGAAATTTGACTG 1390
   |||||

QY 68 ACCGCTATCCCATGCTTTTAGTGGATAGATTTATAGATTACAGCCATAAAAAATTTG 127
   |||||
Db 1389 AAGATGTTGCAACGAGTTTCCAAAATAGTGTGTAGAAAGTTTAAACCTATAAAGCTATTC 1330
   |||||

QY 128 TCGCTTATAGATATACATCTTTTAATGAACGCGTGTTTAACGGGCATTTCCCTAATAA 185
   |||||
Db 1329 ATGCTTGTAAACTTATAATGTTAATCGCTTAATGTTGCAGGTGGTGTGCTAGTAA 1272
   |||||

RESULT 11
US-10-185-51435/c
; Sequence 51435, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
```

APPLICANT: DENISE, Sue K.  
APPLICANT: KERR, Richard  
APPLICANT: ROSENFELD, David  
APPLICANT: HOLM, Tom  
APPLICANT: BATES, Stephen  
APPLICANT: FANTIN, Dennis  
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
FILE REFERENCE: MM1100-2  
CURRENT APPLICATION NUMBER: US/10/750,185  
CURRENT FILING DATE: 2003-12-31  
PRIOR APPLICATION NUMBER: US 60/437,482  
PRIOR FILING DATE: 2002-12-31  
NUMBER OF SEQ ID NOS: 64922  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 51435  
LENGTH: 2528  
TYPE: DNA  
ORGANISM: Bovine 19866880723565  
US-10-750-185-51435

Query Match 7.1%; Score 34.2; DB 6; Length 2528;  
Best Local Similarity 50.3%; Pred. No. 1.6;  
Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 18 AAACCTGCAATCTCAATTTTATAGAGCATATCTTACAAATCTTACCTCACCGCTATCC 77  
DB 323 ACACAGGAACCTCACCATTTCACATGCATAGGCTATCAACAAACAACTGCCCATGAAC 264  
QY 78 CATGCTTTAGTGGATAGAAATTATAGATTACAGGCAATTAACCGGCAATTTCCCTTAATA 137  
DB 263 ATAGCTGTGGTTCAGTCAAGGAAGCCAGTCAAGGACAATTAATAATTATAGTTGAACA 204  
QY 138 GAATATCACTTTTAATGAAGACGTTTAAACGGGCAATTTCCCTTAATA 184  
DB 203 CAAACACATCTTTAATAATCACATTTAGCATCAAACTACCCCTAATA 157

RESULT 12  
US-10-750-623-51435/c  
; Sequence 51435, Application US/10750623  
; Publication No. US20050287531A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-1  
; CURRENT APPLICATION NUMBER: US/10/750,623  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 51435  
; LENGTH: 2528  
; TYPE: DNA  
; ORGANISM: Bovine 19866880723565  
US-10-750-623-51435

Query Match 7.1%; Score 34.2; DB 6; Length 2528;  
Best Local Similarity 50.3%; Pred. No. 1.6;  
Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 18 AAACCTGCAATCTCAATTTTATAGAGCATATCTTACAAATCTTACCTCACCGCTATCC 77  
DB 323 ACACAGGAACCTCACCATTTCACATGCATAGGCTATCAACAAACAACTGCCCATGAAC 264  
QY 78 CATGCTTTAGTGGATAGAAATTATAGATTACAGGCAATTAACCGGCAATTTCCCTTAATA 137

Db 263 ATAGCTGTGGTTCAGTCAAGGAAGCCAGTCAGGACAACTATTAAATTATAGTTGAACA 204  
QY 138 GAATATCACTTTTAAATGAAGACGTTTAAACGGGCAATTTCCCTTAATA 184  
Db 203 CAAACACATCTTTAATAATCACATTTAGCATCAAACTACCCCTAATA 157

RESULT 13  
US-11-117-187-192/c  
; Sequence 192, Application US/11117187  
; Publication No. US20050266560A1  
; GENERAL INFORMATION:  
; APPLICANT: PREUSS, DAPHNE  
; APPLICANT: COPENHAVER, GREGORY  
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS  
; FILE REFERENCE: ARCD:309US  
; CURRENT APPLICATION NUMBER: US/11/117,187  
; CURRENT FILING DATE: 2005-04-28  
; PRIOR APPLICATION NUMBER: US/09/531,120  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/125,219  
; PRIOR FILING DATE: 1999-03-18  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 192  
; LENGTH: 134499  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-11-117-187-192

Query Match 7.1%; Score 34; DB 7; Length 134499;  
Best Local Similarity 50.6%; Pred. No. 19;  
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 22 TTGCAATCTCAATTTTATAGAGCATATCTTACAAATTTCTACCTCACGGCTATCCCATG 81  
DB 53102 TTTCTAGGTGTCATTTGTATGTCCTCATAGATAACAAATCTCTAATCTCTGCTAATCAAG 53043  
QY 82 CTTTTAGTGGATAGAAATTATAGATTACAGCCCAATTAATAAATTTGTCGCTTATAAGAT 141  
DB 53042 TTAATTTGGCGTTAAATATCAAAATTTATAGAAATAAACAATTTACAGTTATATCT 52983  
QY 142 ATCACTTTTAAATGAAGACGTTTAAACGGGCAATTTCCCTTAAT 183  
DB 52982 ATTACATATAATTTAGATTCACTCAACATGATGTCACCTTAT 52941

RESULT 14  
US-10-750-185-57788  
; Sequence 57788, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 57788  
; LENGTH: 2382  
; TYPE: DNA  
; ORGANISM: Bovine 19866881295850  
US-10-750-185-57788





GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:12:12 ; Search time 3014 Seconds  
(without alignments)  
9052.706 Million cell updates/sec

Title: US-10-662-126-36  
Perfect score: 480  
Sequence: 1 atgggaacaaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 24

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_in.\*
- 3: gb\_env.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pr.\*
- 9: gb\_ro.\*
- 10: gb\_sts.\*
- 11: gb\_sy.\*
- 12: gb\_un.\*
- 13: gb\_vl.\*
- 14: gb\_htg.\*
- 15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	480	100.0	480	6	BD009867
c	2	480	100.0	13222	1 AE000637
3	294	61.3	853	6	AX788995
4	100	20.8	659	6	AX789107
c	5	86	17.9	14404	1 AE001551
6	61	12.7	480	1	AY725427

ALIGNMENTS

RESULT 1  
BD009867  
LOCUS  
DEFINITION  
Accession  
BD009867  
BD009867.1 GI:18638240  
480 bp DNA linear PAT 31-JAN-2002  
Proteins, in particular membrane proteins, of Helicobacter pylori,  
their preparation and use.

KEYWORDS JP 2001502886-A/8.  
SOURCE Helicobacter pylori  
ORGANISM Helicobacter pylori  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Helicobacteraceae; Helicobacter.  
REFERENCE 1 (bases 1 to 480)  
AUTHORS Knapp,B., Hundt,E. and Schmidt,K.H.  
TITLE Proteins, in particular membrane proteins, of Helicobacter pylori,  
their preparation and use  
JOURNAL .Patent: JP 2001502886-A 8 06-MAR-2001;  
CHIRON BEHRING GMBH & CO  
COMMENT OS Helicobacter pylori  
PN JP 2001502886-A/8  
PD 06-MAR-2001  
PF 25-JUL-1997 JP 1998508651  
PI 26-JUL-1996 DE 196 30 390.7  
PI BERNHARD KNAPP,ERIKA HUNDT,KARL HEINZ SCHMIDT PC  
C12N15/31,C07K14/205,C07K16/12,G01N33/53,A61K31/70,A61K39/106, PC  
A61K39/395  
CC Strandedness: Single;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT source 1..480  
FT /organism='Helicobacter pylori'.  
FEATURES  
source Location/Qualifiers  
1..480  
/organism='Helicobacter pylori'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:210'

Query Match 100.0%; Score 480; DB 6; Length 480;  
Best Local Similarity 100.0%; Pred. No. 3.8e-273;  
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAAACAAAGCCATCAAACTTGCAATCTCAATTTTATAGAGCATATCTTACAAATT 60  
Db 1 ATGGAAACAAAGCCATCAAACTTGCAATCTCAATTTTATAGAGCATATCTTACAAATT 60  
Qy 61 CTACCTCACCGCTATCCCGTCTTTAGTGATAGATTAAGATTATAGAGTTACAGCCCAATAA 120  
Db 61 CTACCTCACCGCTATCCCGTCTTTAGTGATAGATTAAGATTATAGAGTTACAGCCCAATAA 120  
Qy 121 AAAATTGTCGCTTATAAGAAATATCACCTTTAAATGAAGACGTTTAAACGGCATTTCCCT 180  
Db 121 AAAATTGTCGCTTATAAGAAATATCACCTTTAAATGAAGACGTTTAAACGGCATTTCCCT 180  
Qy 181 AATAAGCCCATTTTCCCGGCGTTTTCATCGTAGAGGCGATGGCGCAACCGGAGGTTT 240  
Db 181 AATAAGCCCATTTTCCCGGCGTTTTCATCGTAGAGGCGATGGCGCAACCGGAGGTTT 240  
Qy 241 TTAGCCTTCACTAGCTTGTGGGGTGTGACCTGAAATCGCAAAACAAAATCGTGAT 300  
Db 241 TTAGCCTTCACTAGCTTGTGGGGTGTGACCTGAAATCGCAAAACAAAATCGTGAT 300  
Qy 301 TTCATGACGATTGATAAGGTTAAATTCGCGATCCCTGTAAACCCAGCGGACAGATTAGAA 360  
Db 301 TTCATGACGATTGATAAGGTTAAATTCGCGATCCCTGTAAACCCAGCGGACAGATTAGAA 360  
Qy 361 TACCATTAGAAGTCTTAAAGCATAAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420  
Db 361 TACCATTAGAAGTCTTAAAGCATAAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420  
Qy 421 GTGGATGGCAAAAGTGTGCTGAAGCCGAATTGAAGCCATGATTGCAGAGAGATTAA 480  
Db 421 GTGGATGGCAAAAGTGTGCTGAAGCCGAATTGAAGCCATGATTGCAGAGAGATTAA 480

RESULT 2  
AE000637/c  
LOCUS  
DEFINITION  
Helicobacter pylori 26695 section 115 of 134 of the complete  
genome.

ACCESSION	AB000637	AB000511	
VERSION	AB000637.1	GI:2314536	
KEYWORDS			
SOURCE	Helicobacter pylori 26695		
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
REFERENCE	1 (bases 1 to 13222)		
AUTHORS	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.		
TITLE	The complete genome sequence of the gastric pathogen Helicobacter pylori		
JOURNAL	Nature 388 (6642), 539-547 (1997)		
PUBMED	9252185		
REFERENCE	2 (bases 1 to 13222)		
AUTHORS	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.		
Direct Submission	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
JOURNAL	3 (bases 1 to 13222)		
REFERENCE	White, O.		
AUTHORS	Direct Submission		
TITLE	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1. .13222 /organism="Helicobacter pylori 26695" /mol_type="genomic DNA" /strain="26695" /db_xref="taxon:85962" complement(72. .1343) /gene="HP1366" complement(72. .1343) /gene="HP1366" /note="similar to SP:P23191 GB:X56977 PID:44182 percent identity: 37.14; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type IIS restriction enzyme R protein (MBOIR)" /protein_id="AAB08410.1" /db_xref="GI:2314537" /translation="MPKLEKILLETQLDPSKECLFLANRIKSSDYRGHLHLSQHNRY DQNKITIIQAFNEVGDFLQIRITDMKSPNIIIGEEIYAKVDNICKSEMPQDNS GKQNVQDSRLKKNLFVDMRMLIERNNKNETPTNPYQSNKIYISLTPLAIEFLNA QDLRRNFCVYTALENLLQFGACREVMIELDNHYLDIPEMMFVTFILNIENFTRE IIEYRGYRSLRIQEKELQVDCNPNFNGKLEKRDYHNWKNQAQIIFSLLEQ SVFFETNKERLILKALNEENKQNDKIKRSIKEKALYFEKHGVKKEGPELHIVPLC LARSIEPDLRLKLENIYIDAFNHAKISQTNKHICLIFPKNCDDVLKGLKEQESL YLTIVENLYKLDLQAMLYKINKDLHLSKNG" complement(1359. .2141) /gene="HP1367" complement(1359. .2141) /note="similar to SP:P23192 GB:X56977 PID:44181 percent identity: 59.27; identified by sequence similarity; putative" /codon_start=1 /transl_table=11		
gene		/product="type IIS restriction enzyme M1 protein (mod)" /protein_id="AAB08411.1" /db_xref="GI:2314538" /translation="MILNKIYIEDVFTFLDKLEDKSVDLAIIDPPYNLKIASWDSFKN DEEFTFSYAWIDKMLPKLGTGTFYFNTFNCALFLAYLHHKKHFLNPTITWVKKO GPANAKRNVNHAQESILFYSMHKNTYFNADIRIAVESAEIRIKHAQSGKGLKNNKRW FPNKGLKTDWEITSQRVEKKGKILPKPSPKIKKALIERMIKASSHKNDLILD LPSGQMTSLVAKSLERNFICGSHAEYVHGSLMEFRYNECE" complement(2128. .2991) /gene="HP1368" complement(2128. .2991) /note="similar to SP:P29568 percent identity: 32.96; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type IIS restriction enzyme M2 protein (mod)" /protein_id="AAB08412.1" /db_xref="GI:2314539" /translation="MNINKVFYHSSHTNMNEVPDNSVDLIITSPFYFNKDYAKNGTQD LQSHQHVDELGALEKYEDYLLGLKWLCEYALPKNGKLCINPLMLPKKVLNTH YNRHIFDLHADIQHSILHDLNNMLKNKPMFLDVIYKQANPYKRMFGSYPPRNF YACHTIEFGVFKDGKPKOPTOEKQSOLTOEWEVFTKQIWEIPIPNKNDIAGFK HAALMPAELARRLIRLYSCYGDVLDLPFSSGTTLREAKLLKRNFIGVELYENKPLI EQKLGNDLDFPE" 3240. .4664 /gene="HP1369" /note="This region contains an authentic frame shift and is not the result of a sequencing artifact; similar to SP:P40814 percent identity: 39.74; identified by sequence similarity; putative; type III restriction enzyme M protein (mod)" 4802. .5668 /gene="HP1370" 4802. .5668 /gene="HP1370" /note="similar to GB:L42023 PID:1006312 PID:1221173 PID:1205306 PID:1574609 percent identity: 38.41; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type III restriction enzyme M protein (mod)" /protein_id="AAB08413.1" /db_xref="GI:2314540" /translation="MYKLDENGPKLDPVMAKSGNNPSPYTLNGKWTSPSGTGFWR YSIGTLKMEQNNRIYFNGKNPMKRYLSEVAGRKSGSTFDGSEVGNLNGDAIKQ LFNGKVPNPKPEALLQRILEISTKENDLVDLPFAGSGTCAVAHKKRYIGLEMG EHPESVILPRLLKVIIGFKSGAAKFDGGGVIKVELESYEEILRKIKYENDNDKPLAY DEQSDLVERKNESYTLNIEALEKMGVDIKETLENLWGVGEFFNEKVKFKGNDKEV EILKALKREALIW" 5679. .8585 /gene="HP1371" 5679. .8585 /gene="HP1371" /note="similar to GB:L25415 PID:496157 percent identity: 26.73; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type III restriction enzyme R protein" /protein_id="AAB08414.1" /db_xref="GI:2314541" /translation="MAKKQDLITDNEIFVAQKLAEBELTNEINEPRLERLDFKSPDNN KELLDDYQOALINAFRLMVAFRDFKKNKEFYAQKHYSPAHCDPAKKGNPLLSK HPKVENHCVFENPILNLAFYMATGSGTKTIVI KLVELLSVAIRMGILPKNIMFSA NENLIQOFEKEIEKYNRNDYFKQIDFKSLSVTHKDFYAPKQDSVIKQITLIFYRAD LMNDESKENLLNYKYDNGENYVILDEAHKGNKSKSQALFSLLSLAGFLNPSA TPTESDLITAVYNLSVGWVWKLGYKESVLLKNKSNLNAFKELKOLNDRKEALILAK LLLGLMKRYKTEGYFYDPLMVFTHSVNNKNSDAEIFFKTLARVINDDGSDFLLAK EDLLEELKNPEFLFSDDKDKDYKVKFKEGLKSMDFKGLKEEVFYANNHGLBVIINPK NQGEIAFLNYSKDFGLIRIGDITETWICEKLSKVVKVSKLSFEESYFSQIDKSSI NILVGSRTFDTGWDSTPVSILFLNIGLDDAKKLVKQSGFGVRIESYFNQGRQLAY LDIIDGAIKALKKPNAMLETFLVPTNYTNSLEAILKFKQSENFNMSEKHFILKHRIYD TPIKHALFVPCYRKEQTSILELPESASPFMSKPNKPKDLKEYFNLMSEKHFILKHRIYD	

PKDYMOLKMTQEAHFNKVS.TWHYKDLDMYMSIKGLYPNOKVPKDEFNALDSEKIV  
HFKRLKVRADKEELVKTIOEYKEVAPLDEKTELIIKKIAOGEIDPYTEKHKKQNTKPV  
GAEELKURHYTTLPIKAKNCDMLKVVVKVSESDFLRELLKITETLQNDYDFWAFS  
KIDEHLNLFIFYNNAERKFPDPFIWLEKGGTOIICFIDPKGSKHDIYEHKADA  
QUFKDKIFNPNPNLKI RVILKFKYGDKEVADGYRDYIWKKGLEDFLQLA"  
/gene="HP1372"  
complement(8635..9381)  
/note="HP1372"  
/note="similar to GB:142023 SP:P44475 PID:1004176  
PID:1221943 PID:1204296 percent identity: 33.56;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="rod shape-determining protein (mreC)"  
/protein\_id="AAD08415.1"  
/db\_xref="GI:2314542"  
/translation="MRFYKFKLWLLGILFIYFLDFKGSSTYISDRIKNALMKNLSL  
LDNQVAFQQAQNIKEFKERLILLEALKLENADLERLNSIYPLENPKMTYPTFTWS  
FISLSDTSVSNIPVNLNENKIYGLVSHNQAIAGIAVLEKGRNLNGFLNAHKRCAYSM  
IQGNQVIGFIGNFQELVVDVFI VPSAEINIGDVLTSGLDIGFGAGVFPVGEVSTED  
HYTYSAVLKNALFSGAKLRHVFLSDVKN"  
complement(9385..10428)  
/gene="HP1373"  
complement(9385..10428)  
/note="HP1373"  
/note="similar to SP:P13519 GB:M22055 PID:606191  
PID:1128967 GB:U00096 percent identity: 51.87; identified  
by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="rod shape-determining protein (mreB)"  
/protein\_id="AAD08416.1"  
/db\_xref="GI:2314543"  
/translation="MIFSKLGLFSDHDAIDLTANTIVLVKGGIILNEPSIVAVRM  
GLFDSKAVDILAVGEAKEMLGKTPNSRAIRPMKDCVIADYDITAKMIRYFIEKVKH  
RKTWIRPRIMVCVPGLTSVERNAKBSLTSAGAEVFLIEBPMAAIIAGALPVKPEPQ  
GSLIVDIGGTEIGVISGLGVIISKSIRVAGDKLDQISIVEYIRKFNLLIGERTGEE  
IKIEIGCAIKLDPPTIMEVSGRDQVSGLLHTIELSSDDVFEAIKDQVREISSALARSVL  
BEVKPDLAKDIVONGVVLVTGGGALIKGLDKYLSDMVKLPVYVYDPELLAVAKGTGEAI  
QDLDLLSRVGPSE"  
complement(10471..11811)  
/gene="HP1374"  
complement(10471..11811)  
/note="HP1374"  
/note="similar to GB:118867 SP:P33138 GB:223278 PID:347729  
PID:388977 percent identity: 55.04; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="ATP-dependent protease ATPase subunit (clpX)"  
Query Match 100.0%; Score 480; DB 1; Length 13222;  
Best Local Similarity 100.0%; Pred. No. 3.6e-273;  
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAACAAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
DB 13108 ATGGAACAAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 13049  
QY 61 CTACCTCACCGCTATCCCATCTCAATCTCAATTTTATAGAGCATATCTTACAAATT 120  
DB 13048 CTACCTCACCGCTATCCCATCTCAATCTCAATTTTATAGAGCATATCTTACAAATT 12989  
QY 121 AAAATTGTCGCTTATAAGATATCACTTTTAATGAAGACGTGTTTAAACGGGCATTTCCCT 180  
DB 12988 AAAATTGTCGCTTATAAGATATCACTTTTAATGAAGACGTGTTTAAACGGGCATTTCCCT 12929  
QY 181 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCGATCGGCAACGGAGGGTTT 240  
DB 12928 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCGATCGGCAACGGAGGGTTT 12869  
QY 241 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCCAAAAAATCGTGAT 300

DB 12868 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCAAAATCGTGAT 12809  
QY 301 TTATGACGATTGATAAGGTTAAATTCGCGATCCCTGTAAACCCAGCGACAGATTAGAA 360  
DB 12808 TTATGACGATTGATAAGGTTAAATTCGCGATCCCTGTAAACCCAGCGACAGATTAGAA 12749  
QY 361 TACCATTTAGAGTCTTTAAAGCATTAAGGGCATGATCGGCAAGTGGGTGGCAGCGCTCAA 420  
DB 12748 TACCATTTAGAGTCTTTAAAGCATTAAGGGCATGATCGGCAAGTGGGTGGCAGCGCTCAA 12689  
QY 421 GTGGATCGCAAAAGTGGTCTGCTGAAGCCGAAATTGAAAGCCATTTGAAGCCATTTGACAGAGAGATTAA 480  
DB 12688 GTGGATCGCAAAAGTGGTCTGCTGAAGCCGAAATTGAAAGCCATTTGACAGAGAGATTAA 12629  
RESULT 3  
AX788995 853 bp DNA linear PAT 17-JUL-2003  
LOCUS  
DEFINITION Sequence 1459 from Patent WO02066501.  
ACCESSION AX788995  
VERSION AX788995.1 GI:32955341  
KEYWORDS  
SOURCE Helicobacter pylori  
ORGANISM Helicobacter pylori  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Helicobacteraceae; Helicobacter.  
REFERENCE 1  
AUTHORS Legrain, P., Rain, J.C., Colland, F., de Reuse, H. and Labigne, A.  
TITLE Protein-protein interactions in Helicobacter pylori  
JOURNAL Patent: WO 02066501-A 1459 29-AUG-2002;  
Hybrigenics (FR) ; INSTITUT PASTEUR (FR)  
FEATURES  
Location/Qualifiers  
source 1..853  
/organism="Helicobacter pylori"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:210"  
ORIGIN  
Query Match 61.3%; Score 294; DB 6; Length 853;  
Best Local Similarity 100.0%; Pred. No. 1.4e-162;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAACAAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
DB 560 ATGGAACAAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 619  
QY 61 CTACCTCACCGCTATCCCATCTTGTAGTGAATGAATTTATAGAGTTACAAGCCAAATAA 120  
DB 620 CTACCTCACCGCTATCCCATCTTGTAGTGAATGAATTTATAGAGTTACAAGCCAAATAA 679  
QY 121 AAAATTGTCGCTTATAAGATATCACTTTTAATGAAGACGTGTTTAAACGGGCATTTCCCT 180  
DB 680 AAAATTGTCGCTTATAAGATATCACTTTTAATGAAGACGTGTTTAAACGGGCATTTCCCT 739  
QY 181 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCGATCGGCAACGGAGGGTTT 240  
DB 740 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCGATCGGCAACGGAGGGTTT 799  
QY 241 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCCAAAAAATC 294  
DB 800 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCCAAAAAATC 853  
RESULT 4  
AX789107 659 bp DNA linear PAT 17-JUL-2003  
LOCUS  
DEFINITION Sequence 1571 from Patent WO02066501.  
ACCESSION AX789107  
VERSION AX789107.1 GI:32955397  
KEYWORDS  
SOURCE Helicobacter pylori  
ORGANISM Helicobacter pylori

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.

REFERENCE 1  
AUTHORS Legrain, P., Rain, J.C., Colland, F., de Reuse, H. and Labigne, A.  
TITLE Protein-protein interactions in Helicobacter pylori  
JOURNAL Patent: WO 02066501-A 1571 29-AUG-2002;  
Hybrigenics (FR) ; INSTITUT PASTEUR (FR)  
FEATURES Location/Qualifiers  
source 1..659  
/organism="Helicobacter pylori"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:210"

ORIGIN  
Query Match 20.8%; Score 100; DB 6; Length 659;  
Best Local Similarity 100.0%; Pred. No. 3.1e-47;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
|||||  
Db 560 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 619  
|||||  
QY 61 CTACCTCACCGTATCCATGCTTTTAGTGGATAGAAATTA 100  
|||||  
Db 620 CTACCTCACCGTATCCATGCTTTTAGTGGATAGAAATTA 659  
|||||

RESULT 5  
AB001551/c  
LOCUS Helicobacter pylori, strain J99 section 112 of 132 of the complete genome. BCT 20-JAN-1999  
DEFINITION Helicobacter pylori, strain J99 section 112 of 132 of the complete genome.  
ACCESSION AB001551 AB001439  
VERSION AB001551.1 GI:4155887  
KEYWORDS  
SOURCE Helicobacter pylori J99  
ORGANISM Helicobacter pylori J99  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.  
1 (bases 1 to 14404)  
Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G., Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C., Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E., Vovis, G.F. and Trust, T.J.

TITLE Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori  
JOURNAL Nature 397 (6715), 176-180 (1999)  
PUBMED 9923682  
REFERENCE 2 (bases 1 to 14404)  
AUTHORS King, B.L., Alm, R.A. and Trust, T.J.  
TITLE Direct Submission  
JOURNAL Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney Street, Cambridge, MA 02139, USA

COMMENT Address all correspondence to: hp@arch.us.astra.com or Richard A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith, Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills, Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and Diane E. Taylor are with the University of Alberta Department of Medical Microbiology and Immunology, Edmonton, Alberta, Canada. T6G 2H7 and the Canadian Bacterial Diseases Network. All other authors are with Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Putative identifications, sequence alignments, and name and sequence search capability are available at ARCB's World Wide web site. (URL: http://www.astra-boston.com/hpylori).

FEATURES  
source 1..14404  
/organism="Helicobacter pylori J99"  
/mol\_type="genomic DNA"  
/strain="J99"

gene  
complement (71..1531)  
/gene="dnaB"  
/note="synonym: jhp1280"  
complement (71..1531)  
/gene="dnaB"  
/notes="similar to H. pylori 26695 gene HP1362"  
/codon\_start=1  
/transl\_table=11  
/product="REPLICATIVE DNA HELICASE"  
/protein\_id="AAD06865.1"  
/db\_xref="GI:4155895"  
/translation="MOHLKHLQQLQNIERIVLSGIVLANHKIBEIHVSHPSPDFYPP HGLFFELAKLHEVNCPIDENFTRQKPKDKQISEDDLVAIFAAASPIDNIEAYVEEIK NASIKKFLTLANTIREOALESAQSKSDIILNAVEREVYALINGSTIEGFTGKEVLES TWNLITENQKGLSKVTGPTGVQLDNYTSGFNQSLVILGARPSMGKTSLMNMVLS SALLNDRGVAFVSLEMSAEQLARLSDLTSSINMHDLSEARLDDQWENLAKCFDHL S QKKLFYDKSYVRMDQIRLQKLSQHKELGIAFDYDLQMSGNKATKEHEQIAEI SRELKTLARELEIPIIALVOLNKSLENRRDKPILSDIKOSGGIEQADIVLFIYRGY IYOMRAEDNKIDIKKEGVEEAQELHLKYNEERRIHKQNGSIEEABIIIVAKNRNGAT GTVYTRFNAPFTRYEDPVDVSHLEEQETKFEPTT"  
complement (1542..2936)  
/gene="jhp1281"  
complement (1542..2936)  
/gene="jhp1281"  
/notes="similar to H. pylori 26695 gene HP1363"  
/codon\_start=1  
/transl\_table=11  
/product="putative"  
/protein\_id="AAD06866.1"  
/db\_xref="GI:4155896"  
/translation="MLSVEKGNALDKRVLEWLLSEDIIMENAMALERAVLQNASL AWKEENLEDCELDVDCVGSAPFGLFVDFESLSQKARFKIACDIPSGIDSKGRV KXEPKADTTISMGALSKCLSDKADYIGELKVGHLGVNQIYIEITDPTLLKESDL KLPDRKNAHKGDYGHAVHLGKHSAGLISALSFGSGVSVISQALECEITSNNK PUELVECFENPKLSAFALGMENIPDPFKKMLAPCVLDAGVTHKEVLALEKE VILTPKEPFLSLKSGINISMLDNKLEIARDFSQKYPKVLLKGTANTLTAHQ GRVFINLGSVALAKAGSDVLAGLIVSLLSQNTYPLDAAINASLAHALAGLEFKHY ALTPLDLIEIKRL"  
complement (2939..3973)  
/gene="jhp1282"  
complement (2939..3973)  
/gene="jhp1282"  
/notes="similar to H. pylori 26695 gene HP1364"  
/codon\_start=1  
/transl\_table=11  
/product="putative histidine kinase sensor protein"  
/protein\_id="AAD06867.1"  
/db\_xref="GI:4155897"  
/translation="MASKIASEIIQLHMOTHADYHNALNISRYKDASIALFDSKKR VLYSNIPESADLIKTHKEGFSFGVYLFPSDEFAHLGVAKMLFKNSKPLHPSLY RNWLVVFAFLCVIGSVFLGRILFKPIRNEIIRIDHFLKNTTTHLTPMSALVSL KTLDEDQOHRRIKIAIORMSFYRSLSYLVNQDIERSFALLDKALITKENTLFSFM IDYHLEKSDLVGVWLKAEQDFISLYSNLNNAIKYSVMHGYTHIELTPEFLKVK N LGYEIPDKITKLSVRAREFNSSVLGYGIGLDLVKKVCEKYKMRLEIHSEPSKGSFH ENSFCVQFOQ"  
complement (4107..4748)  
/notes="similar to H. pylori 26695 gene HP1365"  
/codon\_start=1  
/transl\_table=11  
/product="putative TRANSCRIPTIONAL REGULATOR"  
/protein\_id="AAD06868.1"  
/db\_xref="GI:4165002"  
/translation="MQKKIFLEDDYLLSESKEFLEHIGVEYFCAFNGKNAHERLSV ERFNLALLDVQVPEMSLELFKIKNDPLISTPVIFITALQDNATLKNALNGASDYL KQFDFIDELEARIKRFENDDPIEIMPNIYVHQHALNVKKEKELIAPKTAQLLEYL EHGKGIISQALENNLEWQAIDDSLTURTIVIKVLRKLKGNKCIETHKGVGVRFNPL"  
complement (4110..4748)  
/gene="jhp1283"  
4857..6806  
/gene="jhp1284"  
4857..6806

gene  
gene  
CDS

```

/gene="jhp1284"
/codon_start=1
/transl_table=11
/product="TYPE II DNA MODIFICATION ENZYME
(METHYLTRANSFERASE)"
/protein_id="AAD06858.1"
/db_xref="GI:4155889"
/translations="MSEITYKDCPCPELEGILNFEDTSLESKKYNEKVDLIIPDKSI
VKDDTKYTNLNSVLTSSKGTQTKNEILKESNQGVKLDDEILKAFVKFKQGS
DYFHNKAGQGLKEQDLDFEYLFKEMTAFDHRKRLNGINTIKEVALEVALVSEFEN
ELCKIYNLRLVLSHFIVSLDKAKRYDLSKITHPNYPKQWKQMDLNLIAIDL
LGNEFLPDLTYFKLEEEVYKSLFENEINGTLIKSENYQALSKNRYKEAIDCIYI
DPPTNGDPAYIDRFQDSYTWLSIMENRLOLAVDLSPOGNFYHLHLDYRANLYGRMLL
NDLFSKENFENELIWHFRTYQSQIQSNFPRKDSLLWYKNCVNNPFKITYSDNYKD
TVYRWREFIVDNNKIIVYNPADSRFDGYLQSKYKPEKNGDIIATINGYKD
DWDTDIOADDPKQADRELQTLTKRPEKLERIKASSNENSIVCFPAGSGTTCAVA
HKLKRIYIGVMEGHESVILPRKPEKLVIGFKSKALKEFNGGVVIKYVELESYEEIIR
KIXYENDKPLAYEEQVSDLVBRKNESYTLNIEALENMGVDIKETLENLHGVGVFPPN
EKVYKFGNDKEVSLKALKEALI"
6816..9734
/gene="jhp1285"
CDS
6816..9734
/gene="jhp1285"
/notes="similar to H. pylori 26695 gene HP1371"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06859.1"
/db_xref="GI:4155899"
/translations="MAKKQOEVRNNEIFVAQKLAEELNANEINDPLEMLDFKSPDNN
KELLDYQOQALINAFVLVAYFRDFPKSEKFYAFQKHYSFANCGPAPAKKLNHLKS
HFKVNHCVRFENINRLAFYATGSGKTIIVIKVELLSVAMGGLIPKNNIMFESA
NHLIKQFKBEIKYRNKDYKQIDFNKLSKNDKPYRAPKDSLMKEITAFYFRAD
LMSDESKNLNLYKDCWNGENYVILDBAHKGNKTESKROAIFSLSLKGLFNPFS
TTBESDLITAYVNLVSGEMVKLYGKESVLLKKNLNAPKELKOLANDREKBTALLKA
LLLLGQKRYKVEFGYFHDPLMLVFTHSVNMSDAEIPFKTLARVINDDESFSKAK
DOLLEELKPEFLFSGDKOKDKYKIEVFKESLQKVDKMGFKLKEAVFYASNGHLEVIIN
PKNNQIAPKNTNDKVFCLIRIGDITETWIREKLKSVKVSKNLSFKSEYSFQIDKS
SNILVGSRAFTGWDSTPVSILFLNIGLDDDAKLVKQSGFGRVIRIESVKNQORL
AYLEIDEALCKLKPNAEMLEMLFVPTNHASLEALFKFKSGSENGENRGPWRBTKL
EXTRLEHALFVPCYKQNTNALKISQSPKMSKKNPKDKLKHFLHMSKGFLLKHEI
YPKDYALKEMIQTAHFKKVSTWHYKOLDYMISEIKGLYPNQKPKDFEPNALDNEK
IYHFKRVKVKADKEKLIQTIQEVKEHAPLDKETLKKLAQGEIDPIDYTEKKHQDRTP
FKVIDHLNLTPIYDINATERFPDFIFWLQKGGTOICIPIDPKSGKHTDYEHKAD
AYQLPEDKVPNKDDNLKIKVVLKFGYKDDVGERYRDLWIEKGLKLYFFSGFKGLA
"
complement(9784..10530)
/gene="jhp1286"
complement(9784..10530)
/gene="jhp1286"
/notes="similar to H. pylori 26695 gene HP1372"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06860.1"
/db_xref="GI:4155890"
/translations="WRPFKFTLWLGILFIYFLDIKGSSTYSIDRVKSALMSAKNSL
LNVQAYFPOAQNIKEFKERLILRALKLENADLKERLNSITPLENPKMTYTPFTNIS
FINLEDTHSVLSNPLNVEENKIYGLVSNQAIAGIAVLEKGRINGFLNAHKRCAYSM
IQONQVLGFTGNFKQELVDFIVFSAEINIQQVLTSLGLDGTFGAGVFPVGEVSSVD
HYTKSAVLKNAPLSEAKLRHVLSGVKN"
complement(10534..11577)
/gene="mreB"
/notes="synonym: jhp1287"
complement(10534..11577)
/gene="mreB"
/notes="similar to H. pylori 26695 gene HP1373"
/codon_start=1
/transl_table=11
/product="ROD SHAPE-DTERMINING PROTEIN"
/protein_id="AAD06861.1"
/db_xref="GI:4155891"

```

```

/translations="WIFSKLIGLFSHDIAIDLGTANTIVLVKGGIINERISIVAVRM
GLPDSKAYDILAVGSEAKEMLGKTPNSIRAIRPMKQDVIADYDITAKMIRYIEKAHK
RKTWIRPRIMVCVPYIGLTVSVERNAKESALSAGAREVFLIEFPMAAIAGAGLPVKPQ
GSLIVDIGGGTTEIGVLSGLVTSKRSIRVAGDKLDQSIYVIRKKNLLIGERTGEE
IKIBIGCAIKLDPPLTMEVSRQDVSGLLHTIELSSDDVFEAIKQVREISSALRSVL
BEVKPDLVRDIVQNGVVLITGGCALIKGLDKVLSDMVKLPVYVGDPELLAVAKGTGEAI
QDLDLISRVGFS"
complement(11630..12988)
/gene="clpX"
/notes="synonym: jhp1288"

Query Match      17.9%; Score 86; DB 1; Length 14404;
Best Local Similarity 100.0%; Pred. No. 6.3e-39;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 GTGTTTAAACGGGCATTTCCCTTAATAAGCCCATTTTCCCGGGCGTTTTGATCGTAGAGGC 219
      |||||
Db 14125 GTGTTTAAACGGGCATTTCCCTTAATAAGCCCATTTTCCCGGGCGTTTTGATCGTAGAGGC 14066

Qy 220 ATGCGCGCAACGGGAGGGTTTTTTAGC 245
      |||||
Db 14065 ATGCGCGCAACGGGAGGGTTTTTTAGC 14040

RESULT 6
AY725427 480 bp DNA linear BCT 08-JUL-2005
LOCUS Helicobacter pylori (3R)-hydroxymyristoyl-acyl carrier protein
DEFINITION dehydratase (fabz) gene, complete cds.
ACCESSION AY725427 GI:56684724
VERSION AY725427.1
KEYWORDS Helicobacter pylori
SOURCE Helicobacter pylori
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
          Helicobacteriaceae; Helicobacter.
REFERENCE 1 (bases 1 to 480)
AUTHORS Liu, W., Luo, C., Han, C., Peng, S., Yang, Y., Yue, J., Shen, X. and
          Jiang, H.
TITLE A new beta-hydroxyacyl-acyl carrier protein dehydratase (fabz) from
          Helicobacter pylori: Molecular cloning, enzymatic characterization,
          and structural modeling
JOURNAL Biochem. Biophys. Res. Commun. 333 (4), 1078-1086 (2005)
PUBMED 15967411
REFERENCE 2 (bases 1 to 480)
AUTHORS Liu, W., Han, C., Wang, Q., Jiang, H. and Shen, X.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-2004) Drug Discovery and Design Center, Shanghai
          Institute of Materia Medica, Shanghai Institutes for Biological
          Sciences, Chinese Academy of Sciences, Zuchongzhi Road 555,
          Shanghai 201203, China
FEATURES             Location/Qualifiers
     source            1..480
                       /organism="Helicobacter pylori"
                       /mol_type="genomic DNA"
                       /strain="SS1"
                       /db_xref="taxon:210"
     gene              1..480
                       /gene="fabz"
     CDS                1..480
                       /gene="fabz"
                       /function="fatty acid and phosphatidic acid biosynthesis"
                       /note="enzyme"
                       /codon_start=1
                       /transl_table=11
                       /product="(3R)-hydroxymyristoyl-acyl carrier protein
                       dehydratase"
                       /protein_id="AAW22049.1"
                       /db_xref="GI:56684725"
     translation="MEQSHQLSQPFIEHILQILPHRYPMLLVDRLTELQANQKIVA
          YKNITDFVNFHGFNFKNPIFPGLIVVEGMAQGGFLAFTSLMGDFEIAKTIYVFM
          TIDKVFRIPTVTPGRLEHVLKHKGMIVQVGGTAQVDGKVVAEAEAKAMIAERE"
ORIGIN

```

Query Match 12.7%; Score 61; DB 1; Length 480;  
Best Local Similarity 100.0%; Pred. No. 4.8e-24;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 169 GGGCAATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTTCATCGTAGAGGGCATGGCGCAA 228  
Db 169 GGGCAATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTTCATCGTAGAGGGCATGGCGCAA 228  
Qy 229 A 229  
Db 229 A 229

Search completed: January 12, 2006, 04:27:48  
Job time : 3014 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: January 12, 2006, 02:20:57 ; Search time 462 Seconds  
(without alignments)  
6924.356 Million cell updates/sec

Title: US-10-662-126-36  
Perfect score: 480  
Sequence: 1 atggaacaaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 24

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_21.\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*  
14: Geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	8	ACA34928 Prokaryot
2	294	61.3	853	6	ABX66131 Helicobac
3	100	20.8	659	6	ABX66187 Helicobac
4	86	17.9	480	4	AAS53867 Helicobac

ALIGNMENTS

RESULT 1  
ID ACA34928  
AC ACA34928 standard; DNA; 480 BP.  
XX ACA34928;  
XX ACA34928;  
DT 19-JUN-2003 (first entry)  
XX Prokaryotic essential gene #16585.  
DE  
XX

Antisense; ds; prokaryotic essential gene; cell proliferation;  
drug design; gene.  
XX Helicobacter pylori.  
XX WO200277183-A2.  
XX 03-OCT-2002.  
XX 21-MAR-2002; 2002WO-US009107.  
XX 21-MAR-2001; 2001US-00815242.  
XX 06-SEP-2001; 2001US-00948993.  
XX 25-OCT-2001; 2001US-0342923P.  
XX 08-FEB-2002; 2002US-00072851.  
XX 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI: 2003-029926/02.  
XX P-PSDB, ABU31058.  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX Claim 14; SEQ ID NO 22798; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 480 BP; 150 A; 93 C; 107 G; 130 T; 0 U; 0 Other;  
Query Match 100.0%; Score 480; DB 8; Length 480;  
Best Local Similarity 100.0%; Pred. No. 1.6e-247;  
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAACAAAGCCATCAAAACTTCGAATCTCAATTTTATAGACATATCTTACAAATT 60  
DB 1 ATGGAACAAAGCCATCAAAACTTCGAATCTCAATTTTATAGACATATCTTACAAATT 60



QY 61 CTACCTCACCCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAAGCCAATAAA 120  
DB |||||||  
QY 121 AAAATTGTCGCTTATAAGAAATACACTTTTAATAGAGCTGTTTAAACGGGCATTTCCCT 180  
DB |||||||  
QY 181 AATAAGCCCAATTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGGTTT 240  
DB |||||||  
QY 241 TTAGCCTTCACCTAGCTTGTGGGGTTTGACCTCGAAATCGCCAAAACAAAATTCGTGTAT 300  
DB |||||||  
QY 301 TTCATGACGATTGATAAGTTAAATTCGCGATCCCTGTAAACCCAGCGACAGATTAGAA 360  
DB |||||||  
QY 361 TACCATTTAGAAGTCTTAAAGCATAAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420  
DB |||||||  
QY 421 GTGGATGGCAAAAGTGTGCTGCTGAAGCCGAATTGAAAGCCATGATTGCAGAGAGATTAA 480  
DB |||||||

RESULT 2  
ABX66131  
ID ABX66131 standard; DNA; 853 BP.  
XX AC ABX66131;  
XX DT 07-MAY-2003 (first entry)  
XX DE Helicobacter pylori selected interacting domain (SID) DNA #730.  
XX KW Protein-protein interaction; ulcer; selected interacting domain; SID;  
XX KW gene; ds.  
XX OS Helicobacter pylori.  
XX PN WO200266501-A2.  
XX PD 29-AUG-2002.  
XX PF 28-DEC-2001; 2001WO-EP015428.  
XX PR 02-JAN-2001; 2001US-0259302P.  
XX (HYBR-) HYBRIGENICS.  
XX PA (INSP ) INST PASTEUR.  
XX PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;  
XX WPI; 2002-674910/72.  
XX DR P-PSDB; ABUS1387.  
XX PT New complexes of protein-protein interactions in Helicobacter pylori,  
XX useful for identifying modulating compounds for treating or preventing  
XX ulcers in mammals.  
XX PS Claim 7; Page 263; 642pp; English.

XX The invention describes a complex of protein-protein interactions in  
XX Helicobacter pylori selected from 421 complexes given in the  
XX specification. The complex of protein-protein interactions are useful for  
XX screening for agents which modulate the interaction of proteins.  
XX Modulating compounds which binds to a targeted bacterial protein may be  
XX used for treating or preventing ulcers in a human or animal. This  
XX sequence encodes a selected interacting domain (SID), identified via

CC protein-protein interactions  
XX SQ Sequence 853 BP; 258 A; 163 C; 167 G; 265 T; 0 U; 0 Other;  
Query Match 61.3%; Score 294; DB 6; Length 853;  
Best Local Similarity 100.0%; Pred. No. 1.9e-147; Indels 0; Gaps 0;  
Matches 294; Conservative 0; Mismatches 0;  
QY 1 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTTTATAGAGCATATCTTCAAAAT 60  
DB |||||||  
QY 61 CTACCTCACCCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAAGCCAATAAA 120  
DB |||||||  
QY 620 CTACCTCACCCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAAGCCAATAAA 679  
DB |||||||  
QY 121 AAAATTGTCGCTTATAAGAAATACACTTTTAATAGAGCTGTTTAAACGGGCATTTCCCT 180  
DB |||||||  
QY 680 AAAATTGTCGCTTATAAGAAATACACTTTTAATAGAGCTGTTTAAACGGGCATTTCCCT 739  
DB |||||||  
QY 181 AATAAGCCCAATTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGGTTT 240  
DB |||||||  
QY 740 AATAAGCCCAATTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGGTTT 799  
DB |||||||  
QY 241 TTAGCCTTCACCTAGCTTGTGGGGTTTGACCTCGAAATCGCCAAAACAAAATTC 294  
DB |||||||  
800 TTAGCCTTCACCTAGCTTGTGGGGTTTGACCTCGAAATCGCCAAAACAAAATTC 853

RESULT 3  
ABX66187  
ID ABX66187 standard; DNA; 659 BP.  
XX AC ABX66187;  
XX DT 07-MAY-2003 (first entry)  
XX DE Helicobacter pylori selected interacting domain (SID) DNA #786.  
XX KW Protein-protein interaction; ulcer; selected interacting domain; SID;  
XX KW gene; ds.  
XX OS Helicobacter pylori.  
XX PN WO200266501-A2.  
XX PD 29-AUG-2002.  
XX PF 28-DEC-2001; 2001WO-EP015428.  
XX PR 02-JAN-2001; 2001US-0259302P.  
XX (HYBR-) HYBRIGENICS.  
XX PA (INSP ) INST PASTEUR.  
XX PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;  
XX WPI; 2002-674910/72.  
XX DR P-PSDB; ABUS1443.  
XX PT New complexes of protein-protein interactions in Helicobacter pylori,  
XX useful for identifying modulating compounds for treating or preventing  
XX ulcers in mammals.  
XX PS Claim 7; Page 277; 642pp; English.

XX The invention describes a complex of protein-protein interactions in  
XX Helicobacter pylori selected from 421 complexes given in the  
XX specification. The complex of protein-protein interactions are useful for  
XX screening for agents which modulate the interaction of proteins.  
XX Modulating compounds which binds to a targeted bacterial protein may be  
XX used for treating or preventing ulcers in a human or animal. This  
XX sequence encodes a selected interacting domain (SID), identified via

CC protein-protein interactions

SQ Sequence 659 BP; 200 A; 125 C; 123 G; 211 T; 0 U; 0 Other;

Query Match 20.8%; Score 100; DB 6; Length 659;

Best Local Similarity 100.0%; Pred. No. 4.7e-43; Indels 0; Gaps 0;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGACATATCTTACAAATT 60

Db 560 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGACATATCTTACAAATT 619

QY 61 CTACCTCACGCTATCCCATGCTTTTACGTAGAGAAATTA 100

Db 620 CTACCTCACGCTATCCCATGCTTTTACGTAGAGAAATTA 659

RESULT 4

AAS53867

ID AAS53867 standard; DNA; 480 BP.

AC AAS53867;

DT 13-FEB-2002 (first entry)

XX Helicobacter pylori DNA for cellular proliferation protein #321.

DE Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

XX Helicobacter pylori.

OS WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

DR P-PSDB; AUA36008.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 7504; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence encodes an essential prokaryotic

CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 480 BP; 151 A; 94 C; 107 G; 128 T; 0 U; 0 Other;

SQ Query Match 17.9%; Score 86; DB 4; Length 480;

Best Local Similarity 100.0%; Pred. No. 1.6e-35;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GTGTTTAAACGGGCATTTCCCTTAATAAGCCCATTTTCCGGGCGTTTGTATCGTAGAGGCG 219

Db 160 GTGTTTAAACGGGCATTTCCCTTAATAAGCCCATTTTCCGGGCGTTTGTATCGTAGAGGCG 219

QY 220 ATGGCGCAAAACGGGAGGGTTTTTAGC 245

Db 220 ATGGCGCAAAACGGGAGGGTTTTTAGC 245

Search completed: January 12, 2006, 03:37:28

Job time : 462 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:25:53 ; Search time 3072 Seconds  
(without alignments)  
7310.482 Million cell updates/sec

Title: US-10-662-126-36  
Perfect score: 480  
Sequence: 1 atggacaagaagccatcaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO\_NUC  
Gapop\_60.0, Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 24  
Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

- Database : EST:\*
- 1: gb\_est1:\*
  - 2: gb\_est2:\*
  - 3: gb\_est3:\*
  - 4: gb\_hic:\*
  - 5: gb\_est4:\*
  - 6: gb\_est5:\*
  - 7: gb\_est6:\*
  - 8: gb\_est7:\*
  - 9: gb\_gss1:\*
  - 10: gb\_gss2:\*
  - 11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----					

No matches found

Search completed: January 12, 2006, 05:21:44  
Job time : 3072 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:14:52 ; Search time 153 Seconds  
(without alignments)  
5576.664 Million cell updates/sec

Title: US-10-662-126-36  
Perfect score: 480  
Sequence: 1 atggaacaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size: 24

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----					

No matches found

Search completed: January 12, 2006, 04:30:27  
Job time : 153 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:37:34 ; Search time 658 Seconds

(without alignments)

6032.373 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggaacaaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: Oligo NUC

Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 24

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA Main:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	100.0	480	7	US-10-282-122A-22798 Sequence 22798, A
2	480	100.0	480	9	US-10-662-126-36 Sequence 36, Appl
3	86	17.9	480	3	US-09-815-242-7504 Sequence 7504, Ap
4	86	17.9	480	7	US-10-335-977-2213 Sequence 2213, Ap
5	76	15.8	180	7	US-10-335-977-2212 Sequence 2212, Ap

ALIGNMENTS

RESULT 1  
US-10-282-122A-22798  
; Sequence 22798, Application US/10282122A  
; Publication No. US20040029129A1

GENERAL INFORMATION:

- APPLICANT: Wang, Liangsu
- APPLICANT: Zamudio, Carlos
- APPLICANT: Malone, Cheryl
- APPLICANT: Haselbeck, Robert
- APPLICANT: Ohlsen, Kari
- APPLICANT: Zyskind, Judith
- APPLICANT: Wall, Daniel
- APPLICANT: Trawick, John

APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 22798  
LENGTH: 480  
TYPE: DNA  
ORGANISM: Helicobacter pylori  
US-10-282-122A-22798

Query Match 100.0%; Score 480; DB 7; Length 480;

Best Local Similarity 100.0%; Pred. No. 1e-252;

Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
DB 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60  
QY 61 CTACCTCACCGCTATCCCGCTTTTAGTGATAGAGTTATAGAGTTACAGCCAAATAA 120  
DB 61 CTACCTCACCGCTATCCCGCTTTTAGTGATAGAGTTATAGAGTTACAGCCAAATAA 120  
QY 121 AAAATTGTGCTTATAAGATATACCTTTTAATGAAGACGTTTAAACGGGCATTTCCCT 180  
DB 121 AAAATTGTGCTTATAAGATATACCTTTTAATGAAGACGTTTAAACGGGCATTTCCCT 180  
QY 181 AATAAGCCATTTTCCCGCGCTTTGATCGTAGAGGCGATGGCGCAAAACGGGAGGGTTT 240  
DB 181 AATAAGCCATTTTCCCGCGCTTTGATCGTAGAGGCGATGGCGCAAAACGGGAGGGTTT 240  
QY 241 TTAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCGCAAAACAAAATCGTGAT 300  
DB 241 TTAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCGCAAAACAAAATCGTGAT 300  
QY 301 TTCATGACGATTGATAAGGTTTAAATTCGCGATCCCTGTAACCCCGGCGACAGATTAGAA 360  
DB 301 TTCATGACGATTGATAAGGTTTAAATTCGCGATCCCTGTAACCCCGGCGACAGATTAGAA 360  
QY 361 TACCATTAGAAGTCTTAAAGCATAAGGGGATGATCTGGCAAGTGGTGGCACCGCTCAA 420  
DB 361 TACCATTAGAAGTCTTAAAGCATAAGGGGATGATCTGGCAAGTGGTGGCACCGCTCAA 420  
QY 421 GTGGATGGCAAGTGTGCTGTAAGCCGGAATTGAAGCCCATGATTGCAGAGAGATTAA 480  
DB 421 GTGGATGGCAAGTGTGCTGTAAGCCGGAATTGAAGCCCATGATTGCAGAGAGATTAA 480

RESULT 2

US-10-662-126-36  
; Sequence 36, Application US/10662126  
; Publication No. US20050063987A1  
; GENERAL INFORMATION:  
; APPLICANT: Knapp, Bernhard  
; APPLICANT: Hundt, Erika  
; APPLICANT: Schmidt, Karl-Heinz  
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of  
; FILE REFERENCE: Helicobacter Pylori, Their Preparation and Use  
; CURRENT APPLICATION NUMBER: US/10/662,126  
; PRIOR FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: 09/230,158  
; PRIOR FILING DATE: 1999-01-19  
; PRIOR APPLICATION NUMBER: PCT/IB97/00981  
; PRIOR FILING DATE: 1997-07-25  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 36  
; LENGTH: 480  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(480)  
; OTHER INFORMATION: bp mature protein; 17 kD protein from Helicobacter  
; OTHER INFORMATION: pylori  
US-10-662-126-36

Query Match 100.0%; Score 480; DB 9; Length 480;  
Best Local Similarity 100.0%; Pred. No. 1e-252;  
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAAACAAGGCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60  
DB 1 ATGGAAACAAGGCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60  
QY 61 CTACCTCCCGCTATCCCATGCTTTAGTGATAGATATATAGATTACAGCCCAATAA 120  
DB 61 CTACCTCCCGCTATCCCATGCTTTAGTGATAGATATATAGATTACAGCCCAATAA 120  
QY 121 AAAATGTGCTTTATAGAAATATCACTTTTAAATGAAGACGTGTTTAAACGGGCAATTCCT 180  
DB 121 AAAATGTGCTTTATAGAAATATCACTTTTAAATGAAGACGTGTTTAAACGGGCAATTCCT 180  
QY 181 AATAAGCCCATTTTCCCGGCGTTTTCATCGTAGAGGCAATGCGCAACGAGGAGGTTT 240  
DB 181 AATAAGCCCATTTTCCCGGCGTTTTCATCGTAGAGGCAATGCGCAACGAGGAGGTTT 240  
QY 241 TTAGCCCTCCTAGCTGTGGGGGTTTGACCCCTGAATCGCCAAACAAATAATCGTGAT 300  
DB 241 TTAGCCCTCCTAGCTGTGGGGGTTTGACCCCTGAATCGCCAAACAAATAATCGTGAT 300  
QY 301 TTATCATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGCAGAGATTAGAA 360  
DB 301 TTATCATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGCAGAGATTAGAA 360  
QY 361 TACCATTTAGAACTTTTAAAGCATAAGGGGATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420  
DB 361 TACCATTTAGAACTTTTAAAGCATAAGGGGATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420  
QY 421 GTGGATGGCAAGTGTGCTGTAAGCCGGAATTTGAAGCCATGATTGAGAGAGAGATTAA 480  
DB 421 GTGGATGGCAAGTGTGCTGTAAGCCGGAATTTGAAGCCATGATTGAGAGAGAGATTAA 480

RESULT 3

US-09-815-242-7504  
; Sequence 7504, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Onisen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7504  
; LENGTH: 480  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(480)  
US-09-815-242-7504

Query Match 17.9%; Score 86; DB 3; Length 480;  
Best Local Similarity 100.0%; Pred. No. 3.4e-36;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 160 GTGTTTAAACGGGCAATTCCTTAATAAGCCCAATTTCCCGGGGTTTTCATCGTAGAGGC 219  
DB 160 GTGTTTAAACGGGCAATTCCTTAATAAGCCCAATTTCCCGGGGTTTTCATCGTAGAGGC 219  
QY 220 ATGCGCAAAACGGGAGGTTTTCAGC 245  
DB 220 ATGCGCAAAACGGGAGGTTTTCAGC 245

RESULT 4

US-10-335-977-2213  
; Sequence 2213, Application US/10335977  
; Publication No. US20040052799A1  
; GENERAL INFORMATION:  
; APPLICANT: DOUGLAS SMITH et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; RELATING TO HELICOBACTER PYLORI FOR  
; DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 10031  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows NT 4.0  
; SOFTWARE: UNIX

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...480
; SEQUENCE DESCRIPTION: SEQ ID NO: 2213:
US-10-335-977-2213

Query Match      17.9%; Score 86; DB 7; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GTGTTTAACGGGCATTCCCTAATAAGCCCATTTCCCGGGCGTTTGTATCGTAGAGGC 219
Db 160 GTGTTTAACGGGCATTCCCTAATAAGCCCATTTCCCGGGCGTTTGTATCGTAGAGGC 219

QY 220 ATGGCGCAACGGGAGGTTTTTAGC 245
Db 220 ATGGCGCAACGGGAGGTTTTTAGC 245

RESULT 5
US-10-335-977-2212
; Sequence 2212, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
```

```
;
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2212:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...180
; SEQUENCE DESCRIPTION: SEQ ID NO: 2212:
US-10-335-977-2212

Query Match      15.8%; Score 76; DB 7; Length 180;
Best Local Similarity 100.0%; Pred. No. 1e-30;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 GGCATTTCCCTAATAAGCCCATTTCCCGGGCGTTTGTATCGTAGAGGCATGGCGCAA 229
Db 11 GGCATTTCCCTAATAAGCCCATTTCCCGGGCGTTTGTATCGTAGAGGCATGGCGCAA 70

QY 230 CGGAGGGTTTTTAGC 245
Db 71 CGGAGGGTTTTTAGC 86
```

Search completed: January 12, 2006, 05:43:16  
Job time : 659 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:29:48 ; Search time 619 Seconds  
(without alignments)  
627.605 Million cell updates/sec

Title: US-10-662-126-36  
Perfect score: 480  
Sequence: 1 atgggaacaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO-NUC  
Gapop 60.0 , Gapext 60.0

Searched: 6038814 seqs, 404674181 residues

Word size : 24  
Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA\_New.\*  
1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----					

No matches found

Search completed: January 12, 2006, 05:32:10  
Job time : 619 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:19:13 ; Search time 134 Seconds  
(without alignments)  
521.352 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MQSHQNLSQFIEHILQI.....QVDGKVVAEALKAMTAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_21.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001s.\*

5: Geneseq2002s.\*

6: Geneseq2003as.\*

7: Geneseq2003bs.\*

8: Geneseq2004s.\*

9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	830	100.0	159	6	ABU31058 Protein e
2	822	99.0	159	4	Aau36008 Helicobac
3	420	50.6	146	6	ABU26345 Protein e
4	388	46.7	153	9	ABM91923 M. xanthu
5	386	46.5	145	8	ADN27276 Bacterial
6	377	45.4	140	8	ADS42226 Bacterial
7	376	45.3	138	8	ADS30389 Bacterial
8	352.5	42.5	152	8	ADS22800 Bacterial
9	349.5	42.1	162	7	ADF6342 Bacterial
10	348	41.9	172	8	ADS42775 Bacterial
11	347	41.8	139	6	ABU30001 Protein e
12	346	41.7	150	8	ADN17354 Bacterial
13	344.5	41.5	157	8	ADN17554 Bacterial
14	344.5	41.5	172	6	ABU40649 Protein e
15	342	41.2	147	8	ADS28964 Bacterial
16	341.5	41.1	160	7	ABO72472 Pseudomon
17	339.5	40.9	155	6	ABU22686 Protein e
18	339	40.8	138	8	ADS25172 Bacterial
19	339	40.8	138	8	ADS26072 Bacterial
20	339	40.8	138	8	ADS22670 Bacterial
21	339	40.8	138	8	ADS25610 Bacterial
22	338	40.7	146	6	ABU41863 Protein e
23	337	40.6	146	4	Aau36344 Pseudomon
24	337	40.6	146	6	ABU38592 Protein e

25	337	40.6	146	7	ADN49809
26	337	40.6	146	7	ADN49811
27	332	40.0	154	8	ADS24815 Bacterial
28	332	40.0	181	6	ABU50043 Protein e
29	331	39.9	159	8	ADS27899 Bacterial
30	330	39.8	153	6	ABU49492 Protein e
31	329	39.6	144	6	ABU24817 Protein e
32	328	39.5	146	6	ABU39637 Protein e
33	323	38.9	138	8	ADN25760 Bacterial
34	323	38.9	140	8	ADS28418 Bacterial
35	323	38.9	151	4	AAU34452 E. coli c
36	323	38.9	151	6	ABU15025 Protein e
37	323	38.9	151	8	ADN18054 Bacterial
38	322	38.8	141	4	AAU35330 Enterococ
39	322	38.8	141	6	ABU29459 Protein e
40	322	38.8	142	7	ADH86668 Enterococ
41	321	38.7	151	6	ABU31679 Protein e
42	321	38.7	154	6	ABU21632 Protein e
43	320	38.6	141	6	ABU23918 Protein e
44	319.5	38.5	152	6	ABU39542 Protein e
45	319.5	38.5	155	6	ABU19696 Protein e

#### ALIGNMENTS

#### RESULT 1

ABU31058  
ID ABU31058 standard; protein; 159 AA.

XX AC ABU31058;

DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #16585.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Helicobacter pylori.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX XX WPI; 2003-029926/02.

XX DR N-PSDB; ACA34928.

XX PS New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX CC Claim 25; SEQ ID NO 58982; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC of the 6213 antisense sequences given in the specification where expression

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 159 AA;  
Query Match 100.0%; Score 830; DB 6; Length 159;  
Best Local Similarity 100.0%; Pred. No. 4.7e-91; Mismatches 0; Indels 0; Gaps 0;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60  
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60  
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120  
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERD 159  
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERD 159  
RESULT 2  
AAU36008  
ID AAU36008 standard; protein; 159 AA.  
XX  
AC AAU36008;  
XX  
DT 14-FEB-2002 (first entry)  
XX  
DE Helicobacter pylori cellular proliferation protein #321.  
XX  
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;  
KW antibacterial; drug design.  
XX  
OS Helicobacter pylori.  
XX  
PN WO200170955-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US009180.  
XX  
PR 21-MAR-2000; 2000US-0191078P.  
PR 23-MAY-2000; 2000US-0206848P.  
PR 26-MAY-2000; 2000US-0207272P.  
PR 23-OCT-2000; 2000US-0242578P.  
PR 27-NOV-2000; 2000US-0253625P.  
PR 22-DEC-2000; 2000US-0257931P.  
PR 16-FEB-2001; 2001US-0269308P.  
XX

PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
DR WPI; 2001-611495/70.  
DR N-PSDB; AAS53867.  
XX  
XX New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.  
XX  
XX Example 3; SEQ ID NO 11601; 511pp; English.  
XX  
CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes,  
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*  
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 159 AA;  
Query Match 99.0%; Score 822; DB 4; Length 159;  
Best Local Similarity 98.7%; Pred. No. 4.3e-90; Mismatches 2; Indels 0; Gaps 0;  
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60  
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60  
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120  
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERD 159  
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERE 159  
RESULT 3  
ABU26345  
ID ABU26345 standard; protein; 146 AA.  
XX  
AC ABU26345;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #11872.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Campylobacter jejuni.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR



















GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

QM protein - protein search, using sw model

Run on: January 10, 2006, 13:26:59 ; Search time 40 Seconds  
(without alignments)  
382.462 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHQNLQSQFFIEHILQI.....QVDGKVAEAEELKAMIAERD 159

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830	100.0	159	2 H64691	(3R)-hydroxymyrist
2	822	99.0	159	2 H71826	(3R)-hydroxymyrist
3	420	50.6	146	2 H81445	(3R)-hydroxymyrist
4	346	41.7	150	2 A70305	(3R)-hydroxymyrist
5	345.5	41.6	173	2 AB3356	(3r)-hydroxymyrist
6	339	40.8	154	2 AG2746	hypothetical prote
7	339	40.8	163	2 F97527	fabz protein (U516
8	337	40.6	146	2 E83190	(3R)-hydroxymyrist
9	332	40.0	181	2 AF0129	(3R)-hydroxymyrist
10	332	40.0	201	2 S35968	(3R)-hydroxymyrist
11	331	39.9	159	2 C87486	hypothetical prote
12	330	39.8	153	2 C82101	(3R)-hydroxymyrist
13	324	39.0	151	2 F90651	hypothetical prote
14	323	38.9	140	2 G84116	hydroxymyristoyl-(
15	323	38.9	151	2 D64742	(3R)-hydroxymyrist
16	323	38.9	151	2 F85502	(3R)-hydroxymyrist
17	320	38.6	141	2 C97338	hydroxymyristoyl-(
18	318	38.3	137	2 B72335	(3R)-hydroxymyrist
19	317	38.2	151	2 AE0530	(3R)-hydroxymyrist
20	316.5	38.1	145	2 A97701	hypothetical prote
21	313	37.7	144	2 AD1390	hydroxymyristoyl-(
22	313	37.7	144	2 AF1765	hydroxymyristoyl-(
23	311.5	37.5	145	2 B71708	(3R)-hydroxymyrist
24	309.5	37.3	140	2 H97919	(3R)-hydroxymyrist
25	309.5	37.3	148	2 G64180	(3R)-hydroxymyrist
26	308.5	37.2	140	2 B95049	hypothetical prote
27	308.5	37.2	144	2 AE6722	hypothetical prote
28	305.5	36.8	166	2 S75045	(3R)-hydroxymyrist
29	298.5	36.0	151	2 E86695	hypothetical prote

30	298	35.9	132	2 D70065	(3R)-hydroxymyrist
31	297.5	35.8	171	2 AH2089	(3R)-hydroxymyrist
32	288	34.7	134	2 C37083	probable (3R)-hydr
33	284	34.2	219	2 T50020	(3R)-hydroxymyrist
34	281.5	33.9	145	2 C84610	probable beta-hydr
35	277	33.4	146	2 H90002	hypothetical prote
36	276	33.3	149	2 G82000	(3R)-hydroxymyrist
37	275	33.1	149	2 D81228	(3R)-hydroxymyrist
38	269	32.4	160	2 C82731	(3r)-hydroxymyrist
39	267.5	32.2	153	2 H86571	myristoyl-acyl car
40	267.5	32.2	153	2 D72051	(3R)-hydroxymyrist
41	266.5	32.1	153	2 C71502	probable (3R)-hydr
42	265.5	32.0	153	2 H81661	(3R)-hydroxymyrist
43	252	30.4	251	2 D75439	(3R)-hydroxymyrist
44	164	19.8	156	2 B83881	(3R)-hydroxymyrist
45	148	17.8	159	2 AC2773	hypothetical prote

ALIGNMENTS

RESULT 1

H64691

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pylori  
C;Species: Helicobacter pylori  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C;Accession: H64691

R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N  
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: H64691

A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA

A;Residues: 1-159 <TOM>

A;Cross-references: UNIPROT:O25928; UNIPARC:UPI000012A4AD; GB:AE000637; GB:AE000511; NID:

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 100.0%; Score 830; DB 2; Length 159;

Best Local Similarity 100.0%; Pred. No. 7e-73;

Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDEVNGHFP 60

Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDEVNGHFP 60

Qy 61 NKPTFPGLIVEGMAQTGGFLAFTSLMGDFPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120

Db 61 NKPTFPGLIVEGMAQTGGFLAFTSLMGDFPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120

Qy 121 YHLEVLKHKGMIMQVGCTAQVDGKVAEAEELKAMIAERD 159

Db 121 YHLEVLKHKGMIMQVGCTAQVDGKVAEAEELKAMIAERD 159

RESULT 2

H71826

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pylori  
C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C;Accession: H71826

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.  
A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: H71826

A;Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-159 <ARN>  
A:Cross-references: UNIPROT:Q9ZJL6; UNIPARC:UPI0000124AC; GB:AE001551; GB:AE001439; NID:  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: fabZ  
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C:Keywords: carbon-oxygen lyase; hydro-lyase; hydrolase

Query Match 99.0%; Score 822; DB 2; Length 159;  
Best Local Similarity 98.7%; Pred. No. 4.1e-72;  
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMLLDRIELQANKKIYAYKNITFNEDVFNHGHP 60  
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMLLDRIELQANKKIYAYKNITFNEDVFNHGHP 60

QY 61 NKPIPPGVLIIVEGMAQTGFLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLE 120  
DB 61 NKPIPPGVLIIVEGMAQTGFLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLE 120

QY 121 YHLEVLKHKGMIVQVGTAGQVDGKVAEAEKAMIAERD 159  
DB 121 YHLEVLKHKGMIVQVGTAGQVDGKVAEAEKAMIAERE 159

RESULT 3  
H81445  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) Cj0273 [imported]  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: H81445  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
R.P.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: H81445  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-146 <PAR>  
A:Cross-references: UNIPROT:Q9PIM2; UNIPARC:UPI0000124A2; GB:AL1139074; GB:AL111168; NID  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: fabZ; Cj0273  
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 50.6%; Score 420; DB 2; Length 146;  
Best Local Similarity 56.8%; Pred. No. 2.4e-33;  
Matches 83; Conservative 24; Mismatches 35; Indels 4; Gaps 2;

QY 14 IEHILQILPHRYPMLLDRIELQANKKIYAYKNITFNEDVFNHGHPNKPDPGVLIVEG 73  
DB 4 VMOIQEILPHRYPFLLDVKITELKVEWLGXKNISIDHVFHMGHPGHPPIYGVLLIEG 63

QY 74 MAQTGFLAFTSLWG-FDPIAKTKIVYFMTIDKVKRIPVTPGDRLEHYHLEVLKHKGM 132  
DB 64 MAQTGGVLIVEGMEDKVDPP---KSKVYFTGIDGAKFRNVRPGDRLDYEMSVYKNGNM 120

QY 133 WQVGGTAQVDGKVAEAEKAMIAER 158  
DB 121 WIFGQAQFVDGNLVAEAEKAMIVDK 146

RESULT 4  
A70305  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: A70305  
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: A70305  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-150 <AQF>  
A:Cross-references: UNIPROT:O66468; UNIPARC:UPI0000056250; GB:AE000671; NID:g2982793; PII  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: fabZ  
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 41.7%; Score 346; DB 2; Length 150;  
Best Local Similarity 50.7%; Pred. No. 3.5e-26;  
Matches 71; Conservative 26; Mismatches 39; Indels 4; Gaps 2;

QY 14 IEHILQILPHRYPMLLDRIELQANKKIYAYKNITFNEDVFNHGHPNKPDPGVLIVEG 73  
DB 3 IQEIMEILPHRYPILLVDKILEIEGKRIIIGLKNVSNPEVFGHPGFPPLPFGVYILEA 62

QY 74 MAQTGFLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLEHYHLEVLKHKGM 133  
DB 63 MAQVGGILMIKSL---NLEIGKYAVV-FAGIDARFKKPVTPGDOILELEVISLKKALS 118

QY 134 QVGGTAQVDGKVAEAEKLA 153  
DB 119 KMKGVAKVDGEVVAQATLMA 138

RESULT 5  
AB3356  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) [imported] - Bruce  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 17-May-2002  
C:Accession: AB3356  
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I  
.; Mazur, M.; Goltsman, B.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AB3356  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-173 <KUR>  
A:Cross-references: UNIPARC:UPI0000057E2A; GB:AE008917; PIDN:AA152013.1; PID:g17982777; C  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0832  
A:Map position: 1  
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 41.6%; Score 345.5; DB 2; Length 173;  
Best Local Similarity 44.7%; Pred. No. 4.7e-26;  
Matches 71; Conservative 24; Mismatches 59; Indels 5; Gaps 1;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMLLDRIELQANKKIYAYKNITFNEDVFNHGHP 60  
DB 17 MSDDNQTKLEAADIQALLAVLPHRYPFLLDRIIDVDIGDVSATGINKVNTINEPHTGHP 76

QY 61 NKPIPPGVLIIVEGMAQTGFLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLE 120  
DB 77 ENPIMPGLIVVMAQTAGAISL-----LQRTGTPGVVYFMTIDSAKFRPVPVPGDRLL 131

QY 121 YHLEVLKHKGMIVQVGTAGQVDGKVAEAEKAMIAERD 159  
DB 132 LYVKIKORANISKVECVAEVDGKVAEAEVAMISVAD 170

RESULT 6  
AG2746

hypothetical protein fabZ [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AG2746  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AG2746  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-154 <KUR>  
A:CROSS-references: UNIPROT:Q8UFL4; UNIPARC:UPI000012A49B; GB:AE008688; PIDN:AAL442389.1;  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: fabZ  
A:Map position: circular chromosome  
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
Query Match 40.8%; Score 339; DB 2; Length 154;  
Best Local Similarity 44.9%; Pred. No. 1.7e-25;  
Matches 66; Conservative 29; Mismatches 36; Indels 16; Gaps 2;  
QY 14 IEHILQILPHRYPMMLVDRIELQANKKIVAYKNITFEDVFNHGFNPKPIPPGVLIIVEG 73  
DB 14 ILEVKKLLPHRYPFLLIDKIIIDGSSAIGIKVTVNEPHFTGHFPDRPIMPGLIVEA 73  
QY 74 MAQTGGFLAFTSLMGDFDEIATK-----IVYPMITDKVKFRIPVTPGDRLEHYHLEVLKH 128  
DB 74 MAQTAGAIC-----ARNQGGGHLVYFMTIDNARFRPVPVGGDRLETHVVKQRQ 122  
QY 129 KGMIVQVGTAQVDGKVVAAELKAMI 155  
DB 123 RGNVFKFHCFAKVEGALVAEADVGM 149  
RESULT 7  
F97527  
fabZ protein (U51683) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
C:Accession: F97527  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: F97527  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-163 <KUR>  
A:CROSS-references: UNIPARC:UPI000016438F; GB:AE007869; PIDN:AAK87175.1; PID:g15156449;  
C:Genetics:  
A:Gene: AGR\_C2558  
A:Map position: circular chromosome  
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
Query Match 40.8%; Score 339; DB 2; Length 163;  
Best Local Similarity 44.9%; Pred. No. 1.8e-25;  
Matches 66; Conservative 29; Mismatches 36; Indels 16; Gaps 2;  
QY 14 IEHILQILPHRYPMMLVDRIELQANKKIVAYKNITFEDVFNHGFNPKPIPPGVLIIVEG 73  
DB 23 ILEVKKLLPHRYPFLLIDKIIIDGSSAIGIKVTVNEPHFTGHFPDRPIMPGLIVEA 82  
QY 74 MAQTGGFLAFTSLMGDFDEIATK-----IVYPMITDKVKFRIPVTPGDRLEHYHLEVLKH 128  
DB 83 MAQTAGAIC-----ARNQGGGHLVYFMTIDNARFRPVPVGGDRLETHVVKQRQ 131

QY 129 KGMIVQVGTAQVDGKVVAAELKAMI 155  
DB 132 RGNVFKFHCFAKVEGALVAEADVGM 158  
RESULT 8  
E83190  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1-) PA3645 [imported]  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: E83190  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
A:Accession: E83190  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-146 <STO>  
A:CROSS-references: UNIPROT:Q9HX7; UNIPARC:UPI000012A4B1; GB:AE004784; GB:AE004091; NID:  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: fabZ; PA3645  
A:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C:Keywords: carbon-oxygen lyase; hydro-lyase  
Query Match 40.6%; Score 337; DB 2; Length 146;  
Best Local Similarity 49.3%; Pred. No. 2.5e-25;  
Matches 72; Conservative 19; Mismatches 49; Indels 6; Gaps 3;  
QY 14 IEHILQILPHRYPMMLVDRIELQ-ANKKIVAYKNITFEDVFNHGFNPKPIPPGVLIIVE 72  
DB 4 INEIREVLPHPYPLLDVRRVVELDIEGKRIRAYKNVSNIEPFFNGHGFHEHPIMPGLIIE 63  
QY 73 GMAQTGGFLAFTSLMGDFDEIATKIVYFMTIDKVKFRIPVTPGDRLEHYHLEVLKHGMI 132  
DB 64 AMAQAAGILGFKML---DVKPADGTLYYFVGSQKLRFRQPLPGDLQHLAKFISVKRSI 120  
QY 133 WQVGTAQVDGKVVAAELKAMTAER 158  
DB 121 WKFDCHATVDKPKVCASAI--ICAE 144  
RESULT 9  
AF0129  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1-) [import  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AF0129  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AF0129  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-181 <KUR>  
A:CROSS-references: UNIPROT:Q8ZH57; UNIPARC:UPI000012A4C3; GB:AL590842; PIDN:CAC89897.1;  
C:Genetics:  
A:Gene: fabZ  
A:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C:Keywords: carbon-oxygen lyase; hydro-lyase  
Query Match 40.0%; Score 332; DB 2; Length 181;  
Best Local Similarity 46.4%; Pred. No. 9.9e-25;  
Matches 64; Conservative 27; Mismatches 43; Indels 4; Gaps 1;  
QY 14 IEHILQILPHRYPMMLVDRIELQANKKIVAYKNITFEDVFNHGFNPKPIPPGVLIIVEG 73  
DB 11 ILEVKKLLPHRYPFLLIDKIIIDGSSAIGIKVTVNEPHFTGHFPDRPIMPGLIVEA 82







## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:26:04 ; Search time 160 Seconds  
(without alignments)  
701.119 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHNLSQFFIEHLI.....QVDGKVVAELKAMTAERD 159

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	830	100.0	159	1	FABZ_HELPHY
2	822	99.0	159	1	FABZ_HELPHY
3	813	98.0	159	2	Q5G940_HELPHY
4	509.5	61.4	169	1	FABZ_HELHP
5	488	58.8	151	1	FABZ_WOLSU
6	420	50.6	146	1	FABZ_CAMJE
7	420	50.6	146	1	FABZ_CAMJR
8	418	50.4	142	2	Q4HJF5_CAMLA
9	417	50.2	142	2	Q4HGA6_CAMCO
10	403	48.6	146	2	Q4HQA3_CAMUP
11	396.5	47.8	156	2	Q4NCZ4_9DELUT
12	389	46.9	153	1	FABZ_BRAJA
13	384	46.3	154	1	FABZ_RHIME
14	379	45.7	151	1	FABZ_RHOPA
15	375	45.2	150	1	FABZ_GROSL
16	365	44.0	154	1	FABZ_DRSVH
17	361.5	43.6	155	1	FABZ_BARVU
18	358.5	43.2	155	1	FABZ_BARHE
19	357	43.0	142	1	FABZ_CLOTE
20	355.5	42.8	145	1	FABZ_AZOSE
21	352.5	42.5	152	1	FABZ_RHULO
22	351.5	42.3	157	1	FABZ_BRUSU
23	351	42.3	150	1	FABZ_PHOLL
24	349.5	42.1	169	1	FABZ_GLUOX
25	348	41.9	150	1	FABZ_PROPR
26	347.5	41.9	164	1	FABZ_RALSO
27	346	41.7	150	1	FABZ_AQUAE
28	345.5	41.6	157	1	FABZ_BRUBA
29	345.5	41.6	157	1	FABZ_BRUMA
30	344	41.4	151	1	FABZ_ERWCT
31	341	41.1	146	2	Q4ZWR7_PSESSY

32	340	41.0	150	2	Q5E3F0_VIBF1	Q5e3f0 vibrio fiasc
33	339.5	40.9	146	1	FABZ_METCA	Q604u1 methylococc
34	339.5	40.9	155	1	FABZ_BURMA	Q62jd5 burkholderi
35	339.5	40.9	155	1	FABZ_BURPS	Q63t23 burkholderi
36	339	40.8	154	1	FABZ_AGRF5	Q8uf14 agrobacteri
37	338	40.7	146	1	FABZ_PSESM	Q886n2 pseudomonas
38	337	40.6	146	1	FABZ_PSEAE	Q9hxy7 pseudomonas
39	336	40.5	144	1	FABZ_BACLD	Q85e26 bacillus li
40	336	40.5	150	1	FABZ_VIBPA	Q87me8 vibrio para
41	334.5	40.3	467	1	LPXZ_CHLTE	Q8kx0 c lpxc/fabz
42	334	40.2	150	1	FABZ_MANSO	Q65ve3 manheimia
43	333	40.1	157	1	FABZ_SILPO	Q5lsu4 silicibacte
44	332	40.0	150	1	FABZ_DESPS	Q6aj07 desulfotale
45	332	40.0	150	1	FABZ_VIBVU	Q8dbf0 vibrio vuln

#### ALIGNMENTS

#### RESULT 1

ID	FABZ_HELPHY	STANDARD;	PRT;	159 AA.
AC	O25928;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)			
DE	((3R)-hydroxymyristoyl ACP dehydratase)			
GN	Names-fab2; OrderedLocusNames=HP1376;			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;			
OC	Helicobacteraceae; Helicobacter.			
OX	NCBI_TaxID=210;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=26695 / ATCC 700392;			
RX	MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;			
RA	Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,			
RA	Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,			
RA	Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,			
RA	Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,			
RA	Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,			
RA	Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,			
RA	Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,			
RA	Smith H.O., Fraser C.M., Venter J.C.;			
RT	"The complete genome sequence of the gastric pathogen Helicobacter pylori."			
RT	Nature 388:539-547(1997).			
RL	Nature 388:539-547(1997).			
CC	-I- FUNCTION: Involved in saturated fatty acids biosynthesis.			
CC	-I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-I- SIMILARITY: Belongs to the thioester dehydratase family.			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use as long as its content is in no way modified and this statement is not			
CC	removed.			
CC	-----			
CC	EMBL; AS000637; AAD08419.1; -; Genomic_DNA.			
DR	PIR; H64691; H64691.			
DR	TIGR; HPI376; -.			
DR	HMAP; MF_00406; -; 1.			
DR	InterPro; IPR010084; FabZ.			
DR	TIGRFAMs; TIGR01750; fabz2.1.			
KW	Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.			
FT	ACT SITE 58 58 By similarity.			
SQ	SEQUENCE 159 AA; 18196 MW; 715B6FBED872AE32 CRC64;			

Query Match 100.0%; Score 830; DB 1; Length 159;  
Best Local Similarity 100.0%; Pred. No. 1.6e-71;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNTFNEVDVFNHFF 60
Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNTFNEVDVFNHFF 60

Qy 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDEIATKIVVFMIDKVKFRIPVTPGDRLE 120
Db 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDEIATKIVVFMIDKVKFRIPVTPGDRLE 120

Qy 121 YHLEVLKHGKMIWQVGGTAQVDGKVAEAEKAMIAERD 159
Db 121 YHLEVLKHGKMIWQVGGTAQVDGKVAEAEKAMIAERD 159

RESULT 2
FABZ_HELPJ STANDARD; PRT; 159 AA.
AC Q9ZJL6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=FabZ; OrderedLocusNames=JH1290;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., deJonge B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180 (1999).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE001551; AAD06864.1; -; Genomic_DNA.
CC PIR; H71826; H71826.
CC HAMAP; MF 00406; -; 1.
CC InterPro; IPR010084; FabZ.
CC TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT SITE 58 By similarity.
SQ SEQUENCE 159 AA; 18210 MW; C0F16FBD868AE39 CRC64;

Query Match 99.0%; Score 822; DB 1; Length 159;
Best Local Similarity 98.7%; Pred. No. 9.2e-71;
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNTFNEVDVFNHFF 60
Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNTFNEVDVFNHFF 60

Qy 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDEIATKIVVFMIDKVKFRIPVTPGDRLE 120
Db 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDEIATKIVVFMIDKVKFRIPVTPGDRLE 120

Qy 121 YHLEVLKHGKMIWQVGGTAQVDGKVAEAEKAMIAERD 159
Db 121 YHLEVLKHGKMIWQVGGTAQVDGKVAEAEKAMIAERD 159

```

```

RESULT 3
Q5G940_HELPJ PRELIMINARY; PRT; 159 AA.
AC Q5G940;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE ((3R)-hydroxymyristoyl-acyl carrier protein dehydratase.
GN Name=fabZ;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SS1;
RX PubMed=15967411;
RA Liu W., Luo C., Han C., Peng S., Yang Y., Yue J., Shen X., Jiang H.;
RT "A new beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from
RT Helicobacter pylori: Molecular cloning, enzymatic characterization,
RT and structural modeling.";
RL Biochem. Biophys. Res. Commun. 333:1078-1086 (2005).
DR EMBL; AY725427; AAW22049.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016836; F:hydro-lyase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
SQ SEQUENCE 159 AA; 18184 MW; B7F16AAX78315B74 CRC64;

Query Match 98.0%; Score 813; DB 2; Length 159;
Best Local Similarity 97.5%; Pred. No. 6.7e-70;
Matches 155; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNTFNEVDVFNHFF 60
Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNTFNEVDVFNHFF 60

Qy 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDEIATKIVVFMIDKVKFRIPVTPGDRLE 120
Db 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDEIATKIVVFMIDKVKFRIPVTPGDRLE 120

Qy 121 YHLEVLKHGKMIWQVGGTAQVDGKVAEAEKAMIAERD 159
Db 121 YHLEVLKHGKMIWQVGGTAQVDGKVAEAEKAMIAERD 159

RESULT 4
FABZ_HELPJ STANDARD; PRT; 169 AA.
AC Q7U319;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=HH1181;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 31449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906 (2003).

```

[illegible]

```
QY 74 MAQTGGFLAFTSLWG-FDPEIAKTKIVPMTIDKVKFRIPVTPGDRLEHVLKHGMI 132
Db 64 MAQTGGVLAFTSLWG-FDPEIAKTKIVPMTIDKVKFRIPVTPGDRLEHVLKHGMI 120
QY 133 WQVGGTAQVDGKVAEAEKAMIAER 158
Db 121 WIFKGAQFVGNLVAEAEKAMIVDK 146

RESULT 7
FABZ_CAMJR STANDARD; PRT; 146 AA.
AC QSHWJ3;
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE (3R)-hydroxymyristoyl-ACP dehydratase.
GN Name=fabZ; OrderedLocusNames=CJVE0322;
OS Campylobacter jejuni (strain RM1221).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=195099;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15660156; DOI=10.1371/journal.pbio.0030015;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Ravel J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural differences and novel potential virulence mechanisms
RT from the genomes of multiple Campylobacter species.";
RL PLOS Biol. 3:72-85(2005).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; CP000025; AAW34912.1; -; Genomic_DNA.
DR TIGR; CUE0322; -.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 146 AA; 16440 MW; 6FA2656B12563A5 CRC64;

Query Match 50.6%; Score 420; DB 1; Length 146;
Best Local Similarity 56.8%; Pred. No. 3.3e-32;
Matches 83; Conservative 24; Mismatches 35; Indels 4; Gaps 2;

QY 14 IEHILPHRYPMLLVDRILIELOANKKIVAYKNITFNEVDVFNHGHPFNKPIFGVLIVEG 73
Db 4 VMQIQILPHRYPFLVLVDKITELKVEVLGYKNISIDSHVFNHGHPFGHPPIYPGVLILEG 63

QY 74 MAQTGGFLAFTSLWG-FDPEIAKTKIVPMTIDKVKFRIPVTPGDRLEHVLKHGMI 132
Db 64 MAQTGGVLAFTSLWG-FDPEIAKTKIVPMTIDKVKFRIPVTPGDRLEHVLKHGMI 120
QY 133 WQVGGTAQVDGKVAEAEKAMIAER 158
Db 121 WIFKGAQFVGNLVAEAEKAMIVDK 146

RESULT 8
Q4HJP5_CAMLA PRELIMINARY; PRT; 142 AA.
ID Q4HJP5_CAMLA PRELIMINARY; PRT; 142 AA.
AC Q4HGA6;
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-).
GN Name=fabZ; ORFNames=CC00342;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
```

```
Q4HJP5;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-).
GN Name=fabZ; ORFNames=CLA0287;
OS Campylobacter lari RM2100.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306263;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2100;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAPK01000004; EAL54516.1; -; Genomic_DNA.
DR EMBL; AAPK01000004; EAL54516.1; -; Genomic_DNA.
SQ SEQUENCE 142 AA; 16066 MW; D4D1DBA5660BC286 CRC64;

Query Match 50.4%; Score 418; DB 2; Length 142;
Best Local Similarity 56.6%; Pred. No. 5e-32;
Matches 81; Conservative 27; Mismatches 31; Indels 4; Gaps 2;

QY 17 ILQILPHRYPMLLVDRILIELOANKKIVAYKNITFNEVDVFNHGHPFNKPIFGVLIVEGMAQ 76
Db 3 IQKILPHRYPFLVLVDKITELKVEVLGYKNISIDSHVFNHGHPFDHPIYPGVLILEGMAQ 62

QY 77 TGGFLAFTSLWGDPET-AKTKIVYFMTIDKVKFRIPVTPGDRLEHVLKHGMIQOV 135
Db 63 TGGVLAFTSLWGDPET-AKTKIVYFMTIDKVKFRIPVTPGDRLEHVLKHGMIQOV 119

QY 136 GGTAAQVDGKVAEAEKAMIAER 158
Db 120 EGKAFVGDGLVAEAEKAMIVDK 142

RESULT 9
Q4HGA6_CAMCO PRELIMINARY; PRT; 142 AA.
ID Q4HGA6_CAMCO PRELIMINARY; PRT; 142 AA.
AC Q4HGA6;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-).
GN Name=fabZ; ORFNames=CC00342;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
```



```
CC preliminary data.
DR EMBL; AAFJ01000003; EAL56875.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 142 AA; 15937 MW; 2E000F62AC59B479 CRC64;

Query Match 50.2%; Score 417; DB 2; Length 142;
Best Local Similarity 57.3%; Pred. No. 6.2e-32;
Matches 82; Conservative 25; Mismatches 32; Indels 4; Gaps 2;

QY 17 ILQILPHRYPMMLVDRIELQANKKIYAKNITFNEVDVFNHGFNPKPIPPGVLIIVEGMAQ 76
DB 3 IQEILPHRYPFLLVDKITELKVGVEVRGKYNISIDHVFVGHFGHPHPIYGVLLILEGMAQ 62

QY 77 TGGFLAFTSLWGFDPPI-AKTKIVVFMITDKVKFRIPVTPGDRLEVHLVLEKHKGMIVQV 135
DB 63 TGGVLAFESM---DAKVDPKSKVYVFTGIDGAKFRNPVRREGDLRYEMQVVKRGNWIF 119

QY 136 GGTAAVDGKVAEAEELKAMIAER 158
DB 120 KGQAFVDGMLVAEAEELKAMIVDK 142

RESULT 10
Q4HQ3 CAMUP PRELIMINARY; PRT; 146 AA.
AC Q4HQ3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Beta-hydroxyacyl- (Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.1.-).
GN Name=fabZ; ORFNames=CUP1147;
OS Campylobacter upsaliensis RM3195.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306264;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM3195;
RA Fouts D.E., Mongodin E.P., Mandrell R.E., Miller W.G., Raeko D.A.,
RA Jacques R.J., Burkinac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species."
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFJ01000007; EAL52872.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 146 AA; 16591 MW; 8F1CD2BA02891959 CRC64;

Query Match 48.6%; Score 403; DB 2; Length 146;
Best Local Similarity 55.5%; Pred. No. 1.4e-30;
Matches 81; Conservative 24; Mismatches 37; Indels 4; Gaps 2;

QY 14 IEHILQILPHRYPMMLVDRIELQANKKIYAKNITFNEVDVFNHGFNPKPIPPGVLIIVEG 73
DB 4 IMQIQEILPHRYPFLLVDKITELKVGVEVRGKYNISIDHVFVGHFGHPHPIYGVLLILEG 63

QY 74 MAQTGGFLAFTSLWGFDPPI-AKTKIVVFMITDKVKFRIPVTPGDRLEVHLVLEKHKGM 132
DB 64 MAQTGGVLAFESM---DNKVPKPKVYVFTGIDGAKFRNPVRPGDRLDYEMVRMKNRGAL 120

QY 133 WQVGGTAQVDGKVAEAEELKAMIAER 158
DB 121 WIFESKAFVEENLVAEAEELKAMIVDK 146

RESULT 11
QANZC4_9DEL
```

```
QANZC4_9DEL PRELIMINARY; PRT; 156 AA.
AC QANZC4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Beta-hydroxyacyl- (Acyl-carrier-protein) dehydratase FabZ.
GN ORFNames=ADEHDRAFT_3875;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000002; EAL80756.1; -; Genomic_DNA.
SQ SEQUENCE 156 AA; 17343 MW; 38620F1678976741 CRC64;

Query Match 47.8%; Score 396.5; DB 2; Length 156;
Best Local Similarity 53.1%; Pred. No. 6.4e-30;
Matches 76; Conservative 24; Mismatches 40; Indels 3; Gaps 1;

QY 17 ILQILPHRYPMMLVDRIELQANKKIYAKNITFNEVDVFNHGFNPKPIPPGVLIIVEGMAQ 76
DB 16 IQEILPHRYPFLLVDRIEVEFENHRLVAKGVTVNEFFQGHFPAQVPWPGVLIIEALAQ 75

QY 77 TGGFLAFTSLWGFDPPIAKTKIVVFMITDKVKFRIPVTPGDRLEVHLVLEKHKGMIVQV 136
DB 76 AAALLATMSL---KPDEVKDKITVLMGIDGARFRPVPVPGDRLEVEVTKQKGAVMKQT 132

QY 137 GTAQVDGKVAEAEELKAMIAERD 159
DB 133 GVARVDGQVVAEAEFMAMLADRE 155

RESULT 12
FABZ_BRAJA STANDARD; PRT; 153 AA.
AC Q89KQ3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl- (acyl carrier protein) dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=b114851;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
```

RT Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
RL Bradyrhizobium japonicum USDA110.;

CC !- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC !- SUBCELLULAR LOCATION: Cytoplasmic (By similarity)

CC !- SIMILARITY: Belongs to the thioester dehydratase family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC -----  
CC HAMAP; BA000040; BAC50116.1; -; Genomic\_DNA.

DR HAMAP; MF 00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabz; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 58 By similarity.

SQ SEQUENCE 153 AA; 17178 MW; 337FB1B3D5C1FFDC CRC64;

Query Match 46.9%; Score 389; DB 1; Length 153;

Best Local Similarity 49.3%; Pred. No. 3.3e-29; Mismatches 39; Indels 4; Gaps 1;  
Matches 71; Conservative

QY 14 IEHILQLPHRYPMMLVDRIIELOANKKIVAYKNITFNEVDVFNHGFNPKPIPPGVLIVEG 73

DB 14 INAILQLPHRYPMMLVDRIIELOANKKIVAYKNITFNEVDVFNHGFNPKPIPPGVLIVEG 73

QY 74 MAQTGFLAFTSLMGDPDPIAKTKIVYFTIDKVKRIPVTPGDRLEYHLEVLKHGM1W 133

DB 74 MAQTGFLAFTSLMGDPDPIAKTKIVYFTIDKVKRIPVTPGDRLEYHLEVLKHGM1W 133

QY 74 MAQTGAVIGKISVEGTE-----KPRVYFLTIDCKFKPVLPGDTIEYHMRSLGRKTMW 129

DB 74 MAQTGAVIGKISVEGTE-----KPRVYFLTIDCKFKPVLPGDTIEYHMRSLGRKTMW 129

QY 134 QVGGTAQVDGKVVAAELKAMIAE 157

DB 134 WPHGDAKNGVQVAAEDVAGMLTD 153

RESULT 13

FABZ RHIME

ID FABZ RHIME STANDARD; PRT; 154 AA.

AC Q92Q46;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabZ; OrderedLocusNames=R01504; ORFNames=SMC02092;

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI\_TaxID=382;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=1021;

RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;

RA Capela D., Barloy-Hubler P., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kies E., Lelaure V., Masuy D.,

RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsberger U.,

RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021";

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

CC !- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC !- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC !- SIMILARITY: Belongs to the thioester dehydratase family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC -----  
CC EMBL; AL591787; CAC46083.1; -; Genomic\_DNA.

DR HAMAP; MF 00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabz; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 57 By similarity.

SQ SEQUENCE 154 AA; 16994 MW; 6E1BB0EA6F8CDF40 CRC64;

Query Match 46.3%; Score 384; DB 1; Length 154;

Best Local Similarity 50.0%; Pred. No. 1e-28; Mismatches 24; Indels 14; Gaps 2;  
Matches 75; Conservative

QY 14 IEHILQLPHRYPMMLVDRIIELOANKKIVAYKNITFNEVDVFNHGFNPKPIPPGVLIVEG 73

DB 13 IQBILRLPHRYPFLVLDRIIEIDDDNSAIGIKNVYANEPHFTGHFPEKPIPPGVLIVEG 72

QY 74 MAQTGFLAFTSLMGDPDPIAKTKIVYFTIDKVKRIPVTPGDRLEYHLEVLKH 129

DB 73 MAQTGFLAFTSLMGDPDPIAKTKIVYFTIDKVKRIPVTPGDRLEYHLEVLKH 129

QY 130 GMTWQVGGTAQVDGKVVAAELKAMIAERD 159

DB 123 GNIWKPHCDKADKVDGQVLAEDIGMIVSKE 152

RESULT 14

FABZ RHOPA

ID FABZ RHOPA STANDARD; PRT; 151 AA.

AC P61454;

DT 05-JUL-2004 (Rel. 44, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabZ; OrderedLocusNames=RPA2912;

OS Rhodopseudomonas palustris.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Rhodopseudomonas.

OX NCBI\_TaxID=1076;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=CGA009 / ATCC BAA-98;

RX PubMed=14704707; DOI=10.1038/nbt923;

RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,

RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,

RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,

RA Harrison F.H., Gibson J., Harwood C.S.;

RT "Complete genome sequence of the metabolically versatile

RT photosynthetic bacterium Rhodopseudomonas palustris.";

RL Nat. Biotechnol. 22:55-61(2004).

CC !- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC !- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC !- SIMILARITY: Belongs to the thioester dehydratase family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

DR EMBL; BX572602; CAB28353.1; -; Genomic\_DNA.

DR HAMAP; MF 00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabz; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 56 By similarity.

SQ SEQUENCE 151 AA; 17017 MW; 8F9989D2D9B9FC8 CRC64;

Query Match 45.7%; Score 379; DB 1; Length 151;

Best Local Similarity 48.6%; Pred. No. 3e-28;

Matches 70; Conservative 29; Mismatches 41; Indels 4; Gaps 1;

Search completed: January 10, 2006, 13:34:58  
Job time : 161 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:29:09 ; Search time 46 Seconds  
(without alignments)  
285.770 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHNLQSFIEHILQI.....QVDGKVAEELKAMIAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/prodata/1/iaa/5\_COMB.pep.\*

2: /cgn2\_6/prodata/1/iaa/6\_COMB.pep.\*

3: /cgn2\_6/prodata/1/iaa/H\_COMB.pep.\*

4: /cgn2\_6/prodata/1/iaa/PCRTUS\_COMB.pep.\*

5: /cgn2\_6/prodata/1/iaa/RE\_COMB.pep.\*

6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	388	46.7	153	2	US-09-902-540-11122
2	349.5	42.1	162	2	US-09-543-681A-6627
3	341.5	41.1	160	2	US-09-252-991A-21218
4	322	38.8	142	2	US-09-134-000C-4553
5	317	38.2	186	2	US-09-489-039A-12393
6	311.5	37.5	140	2	US-09-583-110-3906
7	311.5	37.5	146	2	US-09-107-433-2892
8	308.5	37.2	140	2	US-09-196-388-2
9	308.5	37.2	160	2	US-10-089-019-18
10	294	35.4	185	2	US-09-540-236-3473
11	283	34.1	156	2	US-09-134-001C-4234
12	278	33.5	166	2	US-10-089-019-10
13	277	33.4	146	2	US-09-339-614-2
14	271	32.7	163	2	US-09-328-352-7242
15	267.5	32.2	153	2	US-09-198-452A-689
16	267.5	32.2	153	2	US-09-438-185A-653
17	233	28.1	117	2	US-09-107-532A-5188
18	193.5	23.3	132	2	US-09-902-540-11447
19	152	18.3	196	2	US-09-543-681A-6762
20	116.5	14.0	149	2	US-09-902-540-11443
21	98	11.8	173	2	US-09-543-681A-7869
22	98	11.8	174	2	US-09-252-991A-25897
23	85.5	10.3	180	2	US-09-248-796A-19898
24	80.5	9.7	489	2	US-09-134-001C-4902
25	79	9.5	467	2	US-09-248-796A-16476
26	77	9.3	316	2	US-09-248-796A-18830
27	75.5	9.1	121	2	US-09-602-777A-322

28	75.5	9.1	121	2	US-09-602-777A-324	Sequence 324, App
29	75	9.0	260	2	US-09-270-767-33502	Sequence 33502, A
30	75	9.0	260	2	US-09-270-767-48719	Sequence 48719, A
31	73.5	8.9	1002	2	US-10-290-579A-187	Sequence 187, App
32	73.5	8.9	1003	2	US-10-290-579A-193	Sequence 193, App
33	73	8.8	195	2	US-09-489-039A-12028	Sequence 12028, A
34	72.5	8.7	311	2	US-09-489-039A-9786	Sequence 9786, Ap
35	72.5	8.7	1002	2	US-10-290-579A-184	Sequence 184, App
36	72.5	8.7	1003	2	US-10-290-579A-188	Sequence 188, App
37	72	8.7	422	2	US-09-248-796A-16762	Sequence 16762, A
38	71.5	8.6	242	2	US-09-173-300-45	Sequence 27781, A
39	71.5	8.6	424	2	US-10-027-450-45	Sequence 45, Appl
40	71.5	8.6	424	2	US-10-104-047-3687	Sequence 3687, Ap
41	71.5	8.6	466	2	US-09-952-060-4	Sequence 4, Appl
42	71.5	8.6	850	2	US-09-952-060-8	Sequence 8, Appl
43	71.5	8.6	875	2	US-07-743-357-10	Sequence 10, Appl
44	71.5	8.6	1003	1	US-10-290-579A-191	Sequence 191, App
45	71.5	8.6	1005	2	US-10-290-579A-191	Sequence 191, App

ALIGNMENTS

RESULT 1

US-09-902-540-11122

; Sequence 11122, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 11122

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Myxococcus xanthus

; US-09-902-540-11122

Query Match	46.7%	Score 388;	DB 2;	Length 153;
Best Local Similarity	50.0%;	Pred. No. 8.5e-38;		
Matches	73;	Conservative	23;	Mismatches 46;
				Indels 4;
				Gaps 2;
Qy	14	IEHILQLPHRYPMMLVDRIIELOANKKIVAYKNITFNEDVFNHFPNKPFPFGLVIVRG	73	
Db	3	IGETLNLPHRYPLLDVDRVEIIPGQKLTAYKNVTINEPFENGHFGHPGVPMPGVLLLEA	62	
Qy	74	MAQTGGPLATSLMGFPDEIAKTVYFMTIDKVKFRIPVTPGDRLEVHLVKHKGMIV	133	
Db	63	LAQATATLAYKS-ENMDP---SRKLTVMGVDGARFRKVPVLPDRLQLEIEVVRHKGAWV	118	
Qy	134	QVGGTAQVDGKVAEELKAMIAERD	159	
Db	119	KTKGLATVDGARVAEGEFLATVVDKD	144	

RESULT 2

US-09-543-681A-6627

; Sequence 6627, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

```

; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6627
; LENGTH: 162
; TYPE: PR1
; ORGANISM: Proteus mirabilis
US-09-543-681A-6627

```

Query Match	42.1%	Score	349.5;	DB 2;	Length	162;		
Best Local Similarity	46.2%;	Pred. No.	3.2e-33;					
Matches	73;	Conservative	27;	Mismatches	47;	Indels	11; Gaps	4;

  

Qy	1	MEOSHQNLSQSOFIEHILQLPLHRYPMLLVDRRIELQAANKKIYAYKNITFNEDVFNGHP	60
Dd	13	MSDNH-TLQ---I EELDLPHRYPELLVDRVDPEEGKFLRAKKNVSNFNEPFFQGHP	67
		: : :                                 : : :	
Qy	61	NKDIFPGVLIVEGAQTGGFLAFTSLMGFOPEIAKTKIYVFMITIDVKFRIPVTPGDRLE	120
Dd	68	GKPIFGVLILEMAQATGILAFKSVGKLEP----GELYTFAAIDGARKEPVLPGDMV	123
		: : :     : : :	
Qy	121	YHLEVLRKHGMIVQVGTAQVDGKVVAEALKAMIAER	158
Dd	124	LEYEFIKERGVARFKGVAKVDGEIACEAEV--MCARR	159
		: : :     : : :           : : :	

```

RESULT 3
US-09-252-991A-21218
; Sequence 21218, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21218
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-21218

```

Query Match	41.1%;	Score 341.5;	DB 2;	Length 160;
Best Local Similarity	47.5%;	Pred. No. 2.8e-32;		
Matches	75;	Conservative 21;	Mismatches 53;	Indels 9; Gaps 4;
Qy	2	EOSHQNLSQOFFTEHILQILPHRYPMLLVDRILIELQ--ANKKIVAYKNITFNEDVFNHGHP	60	
Db	9	EQAPPNNMDD---INEIREYLPHPYFFLLVDRKVEDIEGKRIRAYKNVSINEPFFNHGHP	65	
Qy	61	NKPIFPGLVIEGMAQQGGFLAFTSLMGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE	120	
Db	66	EHPIMPGLVLIENAAQAAGILGFQWL---DVKPADGTLYIFVSGDKLRFQRPVLPGDQLQ	122	
Qy	121	YHLEVLKHGMIVQGGTAQVDGKGVAAEELKAMIAER	158	
Db	123	LHAKFISVGSIMWFDCHATVDDKPVCSAET--ICAER	158	

RESULT 4  
US-09-134-000C-4553  
; Sequence 4553, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

```

; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4553
; LENGTH: 142
; TYPE: prt
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4553

```

Query Match	38.8%;	Score 322;	DB 2;	Length 142;
Best Local Similarity	50.0%;	Pred. No. 4.7e-30;		
Matches	71;	Conservative 19;	Mismatches 46;	Indels 6; Gaps 3;
Qy	14	IEHILQILPHRYPMLLAVDRIELQANKKIWAYKNITENEDVFNGHPNKKDIPFGVLIVEG	73	
Db	6	ITIEIQLPHRYPFLLDSVEEIVPGERVVAKNVTNEQVFGHFPNPNVLPGVLIIIES	65	
Qy	74	MAQTGGFLAFTSLWGDFDEIATKTIYVFMTIDKVKFRIPVTPGDRLEHYHLEVLKHKMIW	133	
Db	66	LAQAGA-VALLSN---PEF-KGKTAYFGGLDRAKFPQKVTPTGDTLILEVELLKVRASAG	119	
Qy	134	QVGSTAQVDGKVVAAELKAMI	155	
Db	120	MGKGVAKVNGKKVAEELTFMI	141	

```

RESULT 5
US-09-489-039A-12393
; Sequence 12393, Application US/09489039A
; Patent NO. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12393
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12393

```

	Query Match	38.2%	Score 317;	DB 2;	Length 186;
	Best Local Similarity	45.7%;	Pred. No. 2.7e-29;		
	Matches 63;	Conservative 26;	Mismatches 45;	Indels 4;	Gaps 1;
Qy	14 IEHLQLPHRYPMLLVDRIELQANKKIWAYKNITENEDVFNHGHPNKPFGPGLIVBG	73			
Db	45 IEKTVELLPHRYPPLLVDRIIDFEEGRFLRAVKVNSVNEPFFQGHFPGKPIPLGVLLIEA	104			
Qy	74 MAOTGGFLATLSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEVHLVHKGMIV	133			
Db	105 MAQATGILAFKSVGKLEP---GELYFFAGIDEARFRKRPVVPQDMIMEVTFEKTTRGLT	160			
Qy	134 QVGTAQVDGKVVAAEEL	151			
Db	161 RFKGVALLVDGKVVCATM	178			

RESULT 6  
US-09-583-110-3906  
; Sequence 3906, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: LYVIN Doucette-Stamm et al.

```
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3906
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3906

Query Match 37.5%; Score 311.5; DB 2; Length 140;
Best Local Similarity 48.6%; Pred. No. 8.1e-29;
Matches 67; Conservative 21; Mismatches 43; Indels 7; Gaps 3;

QY 14 IEHILQILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDEVFNHGHPNKPPIRPGVLIVEG 73
Db 4 IQGIKEALPHRYPMLLVDRVLEV-SEDTIVAIRKNVTINEPFFNGHFPQYPVMPGVLINEA 62

QY 74 MAQTGGFLAFTSLWGFPDEIAKTIVYFMTIDKVKFRIPVTPGDRLEYHLEVLKHKGMIW 133
Db 63 LAQTAGVLELSK-----PE-NKGKLVFYAGMDKVKFKQVVGQDLVMTATFVKRRGTIA 116

QY 134 QVGGAQVDGKVAEAE 151
Db 117 VVEAKAEVDGKLAASGTL 134

RESULT 7
US-09-107-433-2892
; Sequence 2892, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2892:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...146
; SEQUENCE DESCRIPTION: SEQ ID NO: 2892:
US-09-107-433-2892

Query Match 37.5%; Score 311.5; DB 2; Length 146;
Best Local Similarity 48.6%; Pred. No. 8.6e-29;
Matches 67; Conservative 21; Mismatches 43; Indels 7; Gaps 3;

QY 14 IEHILQILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDEVFNHGHPNKPPIRPGVLIVEG 73
Db 10 IQGIKEALPHRYPMLLVDRVLEV-SEDTIVAIRKNVTINEPFFNGHFPQYPVMPGVLINEA 68

QY 74 MAQTGGFLAFTSLWGFPDEIAKTIVYFMTIDKVKFRIPVTPGDRLEYHLEVLKHKGMIW 133
Db 69 LAQTAGVLELSK-----PE-NKGKLVFYAGMDKVKFKQVVGQDLVMTATFVKRRGTIA 122

QY 134 QVGGAQVDGKVAEAE 151
Db 123 VVEAKAEVDGKLAASGTL 140

RESULT 8
US-09-196-388-2
; Sequence 2, Application US/09196388
; Patent No. 6277595
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick
; APPLICANT: Konstantinidis, Alex
; APPLICANT: Russell, Robert
; TITLE OF INVENTION: fabz
; FILE REFERENCE: GM10182
; CURRENT APPLICATION NUMBER: US/09/196,388
; CURRENT FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-196-388-2

Query Match 37.2%; Score 308.5; DB 2; Length 140;
Best Local Similarity 47.8%; Pred. No. 1.8e-28;
Matches 66; Conservative 22; Mismatches 43; Indels 7; Gaps 3;

QY 14 IEHILQILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDEVFNHGHPNKPPIRPGVLIVEG 73
Db 4 IQGIKEALPHRYPMLLVDRVLEV-SEDTIVAIRKNVTINEPFFNGHFPQYPVMPGVLINEA 62

QY 74 MAQTGGFLAFTSLWGFPDEIAKTIVYFMTIDKVKFRIPVTPGDRLEYHLEVLKHKGMIW 133
Db 63 LAQTAGVLELSK-----PE-NKGKLVFYAGMDKVKFKQVVGQDLVMTATFVKRRGTIA 116

QY 134 QVGGAQVDGKVAEAE 151
Db 117 VVEAKAEVDGKLAASGTL 134

RESULT 9
US-10-089-019-18
; Sequence 18, Application US/10089019
; Patent No. 6951729
; GENERAL INFORMATION:
```

```

; APPLICANT: DEWOLF, WALTER E. JR
;
; APPLICANT: KALLENDER, HOWARD
;
; APPLICANT: LONSDALE, JOHN T.
;
; TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID
;
; TITLE OF INVENTION: SYNTHESIS PATHWAY REAGENTS
;
; FILE REFERENCE: GM50068
;
; CURRENT APPLICATION NUMBER: US/10/089,019
;
; CURRENT FILING DATE: 2002-03-25
;
; PRIOR APPLICATION NUMBER: PCT/US00/29451
;
; PRIOR FILING DATE: 2000-10-26
;
; PRIOR APPLICATION NUMBER: 60/161,775
;
; PRIOR FILING DATE: 1999-10-27
;
; NUMBER OF SEQ ID NOS: 37
;
; SOFTWARE: PastSeq for Windows Version 4.0
;
; SEQ ID NO 18
;
; LENGTH: 160
;
; TYPE: PRT
;
; ORGANISM: Streptococcus pneumoniae
;
; US-10-089-019-18

```

	Query Match	37.2%	Score 308.5;	DB 2;	Length 160;
	Best Local Similarity	47.8%;	Pred. No. 2.2e-28;		
	Matches	66;	Conservative	22;	Mismatches 43;
				Indels	7; Gaps 3;
Qy	14	IEHILQLPHRYPMMLVDRIETIEIQANKKIWAYKNITFNEDEVFGHFNKPIFGPGLIVEG	73		
Db	24	IQIGKALPHRYPMMLVDRLVEV-SEETIVAIKKVTINEPFGHFPQYFPMGPGVIMEA	82		
Qy	74	MAOTGGFPLATSLWGFQPEIAKTIKIVYFMTIDKVKFPIPTVPGDRLEYHLEVLKHGMIW	133		
Db	83	LAQTAGVLELSK-----PE-NKGKLVFVAGMDKVFKKKQVVPDQLVMTATFVKRRGTIA	136		
Qy	134	QVGGTAQVDGKVVAAEEL	151		
Db	137	VVEAKAEVDGKLAASGTL	154		

RESULT 10  
 US-09-540-236-3473  
 ; Sequence 3473, Application US/09540236  
 ; Patent No. 6673910  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GARY L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
 ; FILE REFERENCE: 2709.2005-001  
 ; CURRENT APPLICATION NUMBER: US/09/540, 236  
 ; CURRENT FILING DATE: 2000-04-04  
 ; NUMBER OF SEQ ID NOS: 3840  
 ; SEQ ID NO 3473  
 ; LENGTH: 185  
 ; TYPE: PRT  
 ; ORGANISM: M. catarrhalis  
 US-09-540-236-3473

	Query Match	35.4%;	Score 294;	DB 2;	Length 185;
	Best Local Similarity	45.6%;	Pred. No. 1.4e-26;		
	Matches 67;	Conservative 23;	Mismatches 47;	Indels 10;	Gaps 4;
Qy	16	HILQI---LPHRYPMLLVDRILIELQANKKI VAYKNIITFNEDVFNGHPFNKPIFPQVLIVE	72		
Db	39	HYNOIKHYLPHRYPMLIDRVYACKPNEWITGYKXNISNEELFNGHPFNPIFPQVLQVE	98		
Qy	73	GMAQTGFLAFTSLWGEDPEIATKIVY-FMTIDKVKFRIEPTGDRLEHYLEHLKHGKM	131		
Db	99	AMAQLSGILGFISA---GQTADDGYLYLFAGVDKVKFKVVTSGDQLVIRSKVLMNKRD	154		
Qy	132	IMQYGGTAQVDGKVVAAELKAMIAER	158		
Db	155	LYKPECTAHVDGOLTCSAEII--MTARO	179		

RESULT 11

```

US-09-134-001C-4234
; Sequence 4234, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4234
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4234

```

Query Match	34.1%	Score 283;	DB 2;	Length 156;
Best Local Similarity	40.7%;	Pred. No. 2.2e-25;		
Matches	61;	Conservative 21;	Mismatches 62;	Indels 6; Gaps 1;
Qy	8	LQSQFFTEHTLQILPHRYPMLLVDRIELQANKKIYAVKNTNEDVFNHGHPNKPFPFG	67	
Db	12	METIFDYNQIKQILPHRQPELLDKIVEEGKRCVGLKQVSGNEPFGQHFPFYAVWPG	71	
Qy	68	VLIVEGMAQTGGFLAFTSLMGFDPDEIAKTIKIVVFMITDKVKFKRIPVTPFGDRLEYHVLK	127	
Db	72	VLITEALAQTVAMVLS-----EENKGKIALPAGIDKCRFFKQVVGDTLMLEVEITK	125	
Qy	128	HKGMIWVGGTQAQVDGKVAAELKAMIAE	157	
Db	126	IKGPIGKGSAKATVDGOLACSCELTFAIOD	155	

```

RESULT 12
US-10-089-019-10
; Sequence 10, Application US/10089019
; Patent No. 6951729
; GENERAL INFORMATION:
; APPLICANT: DEWOLF, WALTER E. JR
; APPLICANT: KALLENDER, HOWARD
; APPLICANT: LONSDALE, JOHN T.
; TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID
; TITLE OF INVENTION: SYNTHESIS PATHWAY REAGENTS
; FILE REFERENCE: GM50068
; CURRENT APPLICATION NUMBER: US/10/089,019
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: PCT/US00/29451
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/161,775
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-089-019-10

```

	Query Match	33.5%	Score 278;	DB 2;	Length 166;	
	Best Local Similarity	39.1%	Pred. No. 9.2e-25;			
	Matches	59;	Conservative 24;	Mismatches 62;	Indels 6;	Gaps 1
Qy	7	NLQSQFFIEHTLQTLPHRYPMLLVDRIELQANKKIVAYKNITFNEDEVNGHFPNKPFP	66			
	:	:	:	:	:	:
	:	:	:	:	:	:
Dd	20	HMETFDYNQKLIPHRQFLDLIDKKVEYEGRCVAIKQVSGNEPFFQQGHFEVAVMP	79			
	:	:	:	:	:	:
	:	:	:	:	:	:
Qy	67	GVLIIVEGNAQTGGFLAFTSLWGFDPDEIAKTIVIFMTIDKKVKFRIPVPPTGDRLEHYHLEV	126			
	:	:	:	:	:	:
	:	:	:	:	:	:



Db 80 GVLTEALAQTGAVAILNS-----BENKGIKALFAGIDKCRFKRQVVPDGLTLEVEIT 133  
Qy 127 KHGMIWVGSTQAVDQGVKVAEAEKAMIAE 157  
Db 134 KIKGPIKGNKATVDGQLACSCELTFAIQD 164

## RESULT 13

US-09-339-614-2

; Sequence 2, Application US/09339614

; Patent No. 6489139

; GENERAL INFORMATION:

; APPLICANT: Kallender, Howard

; APPLICANT: Van Horn, Stephanie

; APPLICANT: Warren, Patrick V.

; APPLICANT: Lonsdale, John

; TITLE OF INVENTION: fab2

; FILE REFERENCE: GM10228

; CURRENT APPLICATION NUMBER: US/09/339,614

; CURRENT FILING DATE: 1999-06-24

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 146

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-339-614-2

Query Match 33.4%; Score 277; DB 2; Length 146;  
Best Local Similarity 39.3%; Pred. No. 1e-24;  
Matches 59; Conservative 23; Mismatches 62; Indels 6; Gaps 1;

Qy 8 LOSQFFIEHILQILPHRYPMMLVDRIELQANKKIYAYKNITFNEVDVFNHGHPPNKPFP 67

Db 1 METIFDYNQIKQIIPHRQFLLIDKVEVEEGQRCVAIKQVSGNEPFFQGHFPEYAVMPG 60

Qy 68 VLIIVGMAQTGFLAFTSLWGFDPDEIAKTKIVYFMTIDKVKFRIPVTPGDRLEHYHLEVLK 127

Db 61 VLITEALAQTGAVAILNS-----BENKGIKALFAGIDKCRFKRQVVPDGLTLEVEITK 114

Qy 128 HKGMIWVGSTQAVDQGVKVAEAEKAMIAE 157

Db 115 IKGPIKGNKATVDGQLACSCELTFAIQD 144

## RESULT 14

US-09-328-352-7242

; Sequence 7242, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 7242

; LENGTH: 163

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-7242

Query Match 32.7%; Score 271; DB 2; Length 163;  
Best Local Similarity 44.6%; Pred. No. 6e-24;  
Matches 62; Conservative 19; Mismatches 52; Indels 6; Gaps 3;

Qy 14 IEHILQILPHRYPMMLVDRIELQANKKIYAYKNITFNEVDVFNHGHPPNKPFP 73

Db 20 IOTIRQYLPFRYPFLVDVRVTVDN-SIVGYKNVSINEEFLQGHFPEYIPMPGVLIIEA 78

Qy 74 MAQTGGFLAFTSLWGFDPDEIAKTKIVYFMTIDKVKFRIPVTPGDRLEHYHLEVLKHGMI 132

Db 116 DFLISSKGGKAW---AQARVDSQIVTEAELSPALVDKE 151

Db 79 LAQVSGVLGFI---MNNETPKGSLFLFAGAEVRFKKQVVAGDQLVLKSELVMQKRG 134  
Qy 133 WQVGGTAQVDGKVAEAE 151  
Db 135 KYNCTASVDGIVAATAEI 153

## RESULT 15

US-09-198-452A-689

; Sequence 689, Application US/09198452A

; Patent No. 6559294

; GENERAL INFORMATION:

; APPLICANT: Griffais, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prever

; FILE REFERENCE: 9710-003-999

; CURRENT APPLICATION NUMBER: US/09/198,452A

; CURRENT FILING DATE: 1998-11-24

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 689

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-198-452A-689

Query Match 32.2%; Score 267.5; DB 2; Length 153;  
Best Local Similarity 40.9%; Pred. No. 1.4e-23;  
Matches 65; Conservative 25; Mismatches 54; Indels 15; Gaps 7;

Qy 7 NLOSQFFIEHILQILPHRYPMMLVDRII--ELQANKKIYAYKNITFNEVDVFNHGHPPNKPDI 64

Db 2 NQPSVIKRLRELLDLLPHRYPEFLVDKVLSDYDIEA-RSITAKQNVTTINEPFFMGHFPNAPI 60

Qy 65 FPGVLIVGMAQTGFLAFTSLWGFDPDEIAKTKIVYFMTIDKVKFRIPVTPGD--RLLEY 121

Db 61 MPGLVLIIEALAAQAGVLI-----GLVLENDNRKRIALFLGIQAKFPROAVRPGDVLTLQA 115

Qy 122 HLEVLRKHK-GMIWVGSTQAVDQGVKVAEAEKAMIAERD 159

Db 116 DFLISSKGGKAW---AQARVDSQIVTEAELSPALVDKE 151

Search completed: January 10, 2006, 13:36:46

Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:29:45 ; Search time 61 Seconds  
(without alignments)  
1089.096 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHQNLOSQFFIEHILQI.....QVDGKVVAELKXAMIAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/prodata1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/prodata1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/prodata1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/prodata1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830	100.0	159	4	US-10-282-122A-58982
2	830	100.0	159	5	US-10-662-126-37
3	830	100.0	159	3	US-10-965-006-11
4	822	99.0	159	3	US-09-815-242-11601
5	822	99.0	159	4	US-10-335-977-6975
6	420	50.6	146	4	US-10-282-122A-54269
7	386	46.5	145	4	US-10-369-493-9929
8	377	45.4	140	4	US-10-369-493-20656
9	376	45.3	138	4	US-10-369-493-19422
10	352.5	42.5	152	4	US-10-369-493-11833
11	348	41.9	172	4	US-10-369-493-21205
12	347	41.8	139	4	US-10-282-122A-57925
13	346	41.7	150	4	US-10-369-493-7
14	344.5	41.5	157	4	US-10-369-493-207
15	344.5	41.5	172	4	US-10-282-122A-68573
16	342	41.2	147	4	US-10-369-493-17997
17	339.5	40.9	155	4	US-10-282-122A-50610
18	339	40.8	138	4	US-10-369-493-11703
19	339	40.8	138	4	US-10-369-493-14205
20	339	40.8	138	4	US-10-369-493-14643
21	339	40.8	138	4	US-10-369-493-15105
22	338	40.7	146	4	US-10-282-122A-69787
23	337	40.6	146	3	US-09-815-242-11937
24	337	40.6	146	4	US-10-282-122A-66516
25	337	40.6	146	5	US-10-965-006-2
26	337	40.6	146	5	US-10-965-006-4
27	332	40.0	154	4	US-10-369-493-13848

28	332	40.0	181	4	US-10-282-122A-77967	Sequence 77967, A
29	331	39.9	159	4	US-10-369-493-16932	Sequence 16932, A
30	330	39.8	153	4	US-10-282-122A-77416	Sequence 77416, A
31	329	39.6	144	4	US-10-282-122A-52741	Sequence 52741, A
32	328	39.5	146	4	US-10-282-122A-67561	Sequence 67561, A
33	323	38.9	138	4	US-10-369-493-8413	Sequence 8413, Ap
34	323	38.9	140	4	US-10-369-493-17451	Sequence 17451, A
35	323	38.9	151	3	US-09-815-242-10045	Sequence 10045, A
36	323	38.9	151	4	US-10-369-493-707	Sequence 707, App
37	323	38.9	151	4	US-10-282-122A-42949	Sequence 42949, A
38	323	38.9	151	5	US-10-965-006-10	Sequence 10, Appl
39	322	38.8	141	3	US-09-815-242-10923	Sequence 10923, A
40	322	38.8	141	4	US-10-282-122A-57383	Sequence 57383, A
41	321	38.7	151	4	US-10-282-122A-59603	Sequence 59603, A
42	321	38.7	154	4	US-10-282-122A-49556	Sequence 49556, A
43	320	38.6	141	4	US-10-282-122A-51842	Sequence 51842, A
44	319.5	38.5	152	4	US-10-282-122A-67466	Sequence 67466, A
45	319.5	38.5	155	4	US-10-282-122A-47620	Sequence 47620, A

## ALIGNMENTS

### RESULT 1

US-10-282-122A-58982

; Sequence 58982, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 58982

; LENGTH: 159

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-10-282-122A-58982

Query Match 100.0%; Score 830; DB 4; Length 159;  
Best Local Similarity 100.0%; Pred. No. 9.5e-85;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVYKNTITFNEDEVNGHPP 60  
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVYKNTITFNEDEVNGHPP 60  
  
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
  
QY 121 YHLEVLKHGMIWQVGTAQVDGKVAEAEKAMIAERD 159  
DB 121 YHLEVLKHGMIWQVGTAQVDGKVAEAEKAMIAERD 159

RESULT 2  
US-10-662-126-37  
; Sequence 37, Application US/10662126  
; Publication No. US20050063987A1  
; GENERAL INFORMATION:  
; APPLICANT: Knapp, Bernhard  
; APPLICANT: Hundt, Erika  
; APPLICANT: Schmidt, Karl-Heinz  
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of  
; TITLE OF INVENTION: Helicobacter Pylori, Their Preparation and Use  
; FILE REFERENCE: CHIR-0340  
; CURRENT APPLICATION NUMBER: US/10/662,126  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: 09/230,158  
; PRIOR FILING DATE: 1999-01-19  
; PRIOR APPLICATION NUMBER: PCT/IB97/00981  
; PRIOR FILING DATE: 1997-07-25  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 37  
; TYPE: PRT  
; LENGTH: 159  
; ORGANISM: Helicobacter pylori  
US-10-662-126-37

Query Match 100.0%; Score 830; DB 5; Length 159;  
Best Local Similarity 100.0%; Pred. No. 9.5e-85;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVYKNTITFNEDEVNGHPP 60  
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVYKNTITFNEDEVNGHPP 60  
  
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
  
QY 121 YHLEVLKHGMIWQVGTAQVDGKVAEAEKAMIAERD 159  
DB 121 YHLEVLKHGMIWQVGTAQVDGKVAEAEKAMIAERD 159

RESULT 3  
US-10-965-006-11  
; Sequence 11, Application US/10965006  
; Publication No. US2005021462A1  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, ALED  
; APPLICANT: DHARAMSI, AKIL  
; APPLICANT: VEDADI, MASOUD  
; APPLICANT: DOMAGALA, MEGAN  
; APPLICANT: KIMBER, MATTHEW  
; APPLICANT: VALLEE, FRANCOIS  
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA  
; FILE REFERENCE: IPT-257.01  
; CURRENT APPLICATION NUMBER: US/10/965,006

; CURRENT FILING DATE: 2004-10-14  
; PRIOR APPLICATION NUMBER: PCT/CA03/00560  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: 60/373,321  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 3.3  
; SEQ ID NO 11  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-10-965-006-11

Query Match 100.0%; Score 830; DB 5; Length 159;  
Best Local Similarity 100.0%; Pred. No. 9.5e-85;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVYKNTITFNEDEVNGHPP 60  
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVYKNTITFNEDEVNGHPP 60  
  
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
  
QY 121 YHLEVLKHGMIWQVGTAQVDGKVAEAEKAMIAERD 159  
DB 121 YHLEVLKHGMIWQVGTAQVDGKVAEAEKAMIAERD 159

RESULT 4  
US-09-815-242-11601  
; Sequence 11601, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA 011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11601  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-815-242-11601

Query Match 99.0%; Score 822; DB 3; Length 159;  
Best Local Similarity 98.7%; Pred. No. 7.5e-84;  
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEOSHONLQSOFFIEHILQILPHRYPMLLVDRILIELQANKKIYVKNITFEDVFNHGHP 60  
 Db 1 MEOSHONLQSOFFIEHILQILPHRYPMLLVDRILIELQANKKIYVKNITFEDVFNHGHP 60  
 QY 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120  
 Db 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120  
 QY 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159  
 Db 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERE 159

RESULT 5

US-10-335-977-6975  
 ; Sequence 6975, Application US/10335977  
 ; Publication No. US20040052799A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DOUGLAS SMITH et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
 ; RELATING TO HELICOBACTER PYLORI FOR  
 ; DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 10031  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 28 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109-1875  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: IBM PC Compatible  
 ; OPERATING SYSTEM: Windows NT 4.0  
 ; SOFTWARE: UNIX  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/335,977  
 ; FILING DATE: 30-Dec-2002  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/993,002  
 ; FILING DATE: 17-DEC-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mandragouras, Amy E.  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: GTN-018  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617)227-7400  
 ; TELEFAX: (617)742-4214  
 ; INFORMATION FOR SEQ ID NO: 6975:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 159 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Helicobacter pylori  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (8) LOCATION 1...159  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6975:

Query Match 99.0%; Score 822; DB 4; Length 159;  
 Best Local Similarity 98.7%; Pred. No. 7.5e-84;  
 Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 US-10-335-977-6975  
 QY 1 MEOSHONLQSOFFIEHILQILPHRYPMLLVDRILIELQANKKIYVKNITFEDVFNHGHP 60  
 Db 1 MEOSHONLQSOFFIEHILQILPHRYPMLLVDRILIELQANKKIYVKNITFEDVFNHGHP 60  
 QY 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120

Db 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120  
 QY 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159  
 Db 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERE 159

RESULT 6

US-10-282-122A-54269  
 ; Sequence 54269, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 54269  
 ; LENGTH: 146  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; US-10-282-122A-54269

Query Match 50.6%; Score 420; DB 4; Length 146;  
 Best Local Similarity 56.8%; Pred. No. 9e-39;  
 Matches 83; Conservative 24; Mismatches 35; Indels 4; Gaps 2;  
 QY 14 IEHILQILPHRYPMLLVDRILIELQANKKIYVKNITFEDVFNHGHPFNKPIPPGVLIVEG 73  
 Db 4 VMOIQEILPHRYPFLVLDKTELKVEVVGYNKISISDHVFMGHPGHPPIYGVLLIEG 63  
 QY 74 MAQTGGFLAFTSLWG-FDPEIAKTIVYFMTIDKVKFRIPVTPGDRLEHYHLEVLKHGMI 132  
 Db 64 MAQTGGVLAFESMEDKVP---KSKVYFTGIDGAKERNVPRPGDRLDYENSVVKNRNM 120  
 QY 133 WQVGGTAQVDGKVAEAEELKAMIAER 159  
 Db 121 WIFKGOAFVDGNLVAEAEELKAMIVDK 146

RESULT 7  
US-10-369-493-9929  
; Sequence 9929, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 9929  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: magnetite-containing magnetic coccus  
US-10-369-493-9929  
  
Query Match 46.5%; Score 386; DB 4; Length 145;  
Best Local Similarity 50.0%; Pred. No. 5.9e-35;  
Matches 71; Conservative 28; Mismatches 37; Indels 6; Gaps 2;  
  
QY 14 IEHILQILPHRYPMMLVDRIELQANKKIYVKNITFNEVDVFNHGHFNPKPIFPGVLIVEG 73  
DB 6 LQEINSFLPHRYPFLLIDRIVEVEQKRIAIKNSFNEPHEFGHPPDPHPVMPGVLILEA 65  
  
QY 74 MAQTGGFLAFTSLWGF-DPEIAKTKIYVMTIDKVKFRIPVTPGDRLEYHLEVLKHKGM 132  
DB 66 MAQAGALLA-----GYTDPDSVRGQLVYFMAIDKARFKPVLPGHQLNTEMTLLKRRREV 120  
  
QY 133 WOVGGTAQVDGKVVAAELKAM 154  
DB 121 WRFGKAMVDGEVACEAEVWAM 142  
  
RESULT 8  
US-10-369-493-20656  
; Sequence 20656, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 20656  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Rhodospseudomonas palustris  
US-10-369-493-20656  
  
Query Match 45.4%; Score 377; DB 4; Length 140;  
Best Local Similarity 49.0%; Pred. No. 5.7e-34;  
Matches 70; Conservative 28; Mismatches 41; Indels 4; Gaps 1;  
  
QY 14 IEHILQILPHRYPMMLVDRIELQANKKIYVKNITFNEVDVFNHGHFNPKPIFPGVLIVEG 73

Db 2 INTILKTLPHRPFLLIDRVINIREDSGIGIKNVTVNEPAFQCHPFRPYPGVLMIEG 61  
QY 74 MAQTGGFLAFTSLWGFDPPEIAKTKIYVMTIDKVKFRIPVTPGDRLEYHLEVLKHKGM 133  
DB 62 MAQTAGVIGILSVTGT-----KPRAVYFLTIDKCKFRKVPMPGDTVEYHLTRTGRRKTMW 117  
QY 134 QVGGTAQVDGKVVAAELKAMIA 156  
DB 118 WFGGEAKVDGQIVAEADVGAFLA 140  
  
RESULT 9  
US-10-369-493-19422  
; Sequence 19422, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 19422  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-10-369-493-19422  
  
Query Match 45.3%; Score 376; DB 4; Length 138;  
Best Local Similarity 51.4%; Pred. No. 7.3e-34;  
Matches 72; Conservative 20; Mismatches 44; Indels 4; Gaps 2;  
  
QY 14 IEHILQILPHRYPMMLVDRIELQANKKIYVKNITFNEVDVFNHGHFNPKPIFPGVLIVEG 73  
DB 3 IGEILNLLPHRYPFLLVDRIEIPGQKLTAYNKVTINEFFNFGHFPHPVMPGVLILEA 62  
  
QY 74 MAQTGGFLAFTSLWGFDPPEIAKTKIYVMTIDKVKFRIPVTPGDRLEYHLEVLKHKGM 133  
DB 63 LAQATAILAYKS-ENMDP-----SRKLTVMGVDGARFKRPVLPGDRLQLEIEVVRHKGAVM 118  
QY 134 QVGGTAQVDGKVVAAELKA 153  
DB 119 KTKGLATVDGARVAEGEFLA 138  
  
RESULT 10  
US-10-369-493-11833  
; Sequence 11833, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 11833  
; LENGTH: 152  
; TYPE: PRT

```
; ORGANISM: Mesorhizobium loti
US-10-369-493-11833

Query Match      42.5%; Score 352.5; DB 4; Length 152;
Best Local Similarity 46.2%; Pred. No. 3.6e-31;
Matches 66; Conservative 27; Mismatches 45; Indels 5; Gaps 1;

Qy 17 ILQLPHRYPMMLVDRIIELOANKKIYVAYKNITFNEVDVFNHGFNPKPIFGVLIVEGMAQ 76
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 LMKLLPHRYPMMLVDRIIDIDGDDSAIGIKNVTINEPHFQGHPEQVMPGVLIVEAMAQ 74
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 77 TGGFLAFTSLWGFDPETAKTKIVYFMTIDVKVKPRIPVTPGDRLEVHLVLEKHKGMIVQVG 136
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 TAGAICIRSLGA-----SKSLVFLTIDNAKFKPVVPGDQLKIHKVKKIKRGNLKPFA 129
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 137 GTAQVDGKVAEAEAKAMIAERD 159
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 CEALVDGTGAABAEIASAMVMTGD 152
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-10-369-493-21205
; Sequence 21205, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21205
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(172)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-21205

Query Match      41.9%; Score 348; DB 4; Length 172;
Best Local Similarity 48.3%; Pred. No. 1.4e-30;
Matches 70; Conservative 25; Mismatches 44; Indels 6; Gaps 2;

Qy 14 IEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEVDVFNHGFNPKPIFGVLIVEG 73
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 IEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEVDVFNHGFNPKPIFGVLIVEA 90
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 74 MAQTGGFLAFTSLWGFDPETAKTKIVYFMTIDVKVKPRIPVTPGDRLEVHLVLEKHKGMIV 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 91 MAQTGILAFKSV-----GSLAPGKLYFPAIDGARFKRPVLPDGMVLEVEFKERRGVA 146
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 134 QVGGTAQVDGKVAEAEAKAMIAER 158
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 147 RPKGVAKVDGEVACEAEM--MCARR 169
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-10-282-122A-57925
; Sequence 57925, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
```

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57925
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57925

Query Match      41.8%; Score 347; DB 4; Length 139;
Best Local Similarity 53.5%; Pred. No. 1.3e-30;
Matches 76; Conservative 16; Mismatches 44; Indels 6; Gaps 2;

Qy 14 IEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEVDVFNHGFNPKPIFGVLIVEG 73
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 IOBKIEIIPHYPMMLTIDRVEEMVEGERIVAKKNVTINEPFFQGHPEEPVMPGLIVEA 62
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 74 MAQTGGFLAFTSLWGFDPETAKTKIVYFMTIDVKVKPRIPVTPGDRLEVHLVLEKHKGMIV 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 MAQAGA-VALLSLSEQP-----KGTAYFGGLDCAKFRKKVTPGDTLYLEVLKVKASAG 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 134 QVGGTAQVDGKVAEAEAKAMIAER 155
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 IGKGIKAVDGGKVAEAEALTFMI 138
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-10-369-493-7
; Sequence 7, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
```

```
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
US-10-369-493-7

Query Match          41.7%; Score 346; DB 4; Length 150;
Best Local Similarity 50.7%; Pred. No. 1.9e-30;
Matches 71; Conservative 26; Mismatches 39; Indels 4; Gaps 2;

QY 14 IEHILQILPHRYPMLVDRIIEIQANKKIYVKNITFNEDEVNGHFPNKPPIFGVLIVRG 73
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 IOEIMELPHRYPILLVDKILIEBCKRIIGLKNVSNVPEVQGHPPGPPFPGVIVLEA 62
QY 74 MAQTGGFLAFTSLMGDFDPIAKTKIVYFMTIDKVKFRIPVTGDRLEHYHLEVLKHGMIV 133
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 MAQVGGILMIKSL--NLEIGYAVV-FAGIDDARFKDVPYQDQLILELEVISLKKALS 118
QY 134 QVGGTAQVDGKVAEAEILKA 153
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 KMKGVAKVDGEVVAQATLMA 138

RESULT 14
US-10-369-493-207
; Sequence 207, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 207
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(157)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-207

Query Match          41.5%; Score 344.5; DB 4; Length 157;
Best Local Similarity 46.2%; Pred. No. 3e-30;
Matches 73; Conservative 27; Mismatches 47; Indels 11; Gaps 4;

QY 1 MEQSHQNLOSQFFIEHILQILPHRYPMLVDRIIEIQANKKIYVKNITFNEDEVNGHFP 60
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 MSDNH-TLQ----IEILKLLPHRYPFLVDRLDFEKGKFLAVKNVTNVEPFFQGHFP 62
QY 61 NKPPIFGVLIVEGMAQTGFLAFTSLMGDPDEIAKTKIVYFMTIDKVKFRIPVTGDRLE 120
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 GKPIFGVLILEAMAQATGILAFKTVGTLEP----GELYFFAAIDGARFKRPVLPDQMI 118
QY 121 YHLEVLKHGMIVQVGGTAQVDGKVAEAEILKAMIAER 158
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 MEVEFIKERRGVARFGVAKVDGEVACEAEM--WCARR 154

RESULT 15
```

```
US-10-282-122A-68573
; Sequence 68573, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68573
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (12)..(12)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-68573

Query Match          41.5%; Score 344.5; DB 4; Length 172;
Best Local Similarity 45.6%; Pred. No. 3.4e-30;
Matches 72; Conservative 27; Mismatches 48; Indels 11; Gaps 4;

QY 1 MEQSHQNLOSQFFIEHILQILPHRYPMLVDRIIEIQANKKIYVKNITFNEDEVNGHFP 60
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 MSDNH-TLQ----IEILDLPHRYPFLVDRLDFEKGKFLAVKNVSNFNEPFFQGHFP 79
QY 61 NKPPIFGVLIVEGMAQTGFLAFTSLMGDPDEIAKTKIVYFMTIDKVKFRIPVTGDRLE 120
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 GKPIFGVLILEAMAQATGILAFKSVGKLEP----GELYFFAAIDGARFKRPVLPDQMV 135
QY 121 YHLEVLKHGMIVQVGGTAQVDGKVAEAEILKAMIAER 158
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 LEVEFIKERRGVARFGVAKVDGEIACELEM--MCARR 171
```



Search completed: January 10, 2006, 13:37:58  
Job time : 61 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

QM protein - protein search, using sw model

Run on: January 10, 2006, 13:32:15 ; Search time 9 Seconds  
(without alignments)  
150.054 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHQNLSQFFIEHLQI.....QVGGKVVAAELKAMIAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 61141 seqs, 8493638 residues

Total number of hits satisfying chosen parameters: 61141

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New.\*

- 1: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB pep.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB pep.\*
- 4: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB pep.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB pep.\*
- 6: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB pep.\*
- 7: /cgn2\_6/prodata/1/pubpaa/US11\_NEW\_PUB pep.\*
- 8: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	276	33.3	149	6	US-10-467-657-6110
2	276	33.3	149	6	US-10-467-657-7550
3	276	33.3	154	6	US-10-467-657-6130
4	75.5	9.1	121	6	US-10-454-437-322
5	75.5	9.1	121	6	US-10-454-437-324
6	74.5	9.0	485	6	US-10-485-517-136
7	73.5	8.9	393	7	US-11-194-246-309
8	71.5	8.6	564	6	US-10-821-234-1340
9	70.5	8.5	912	7	US-11-042-988-12
10	66	8.0	205	7	US-11-093-746A-26
11	66	8.0	792	7	US-11-103-957-92
12	65.5	7.9	129	7	US-11-052-554A-363
13	65.5	7.9	323	7	US-11-055-822-628
14	65	7.8	1343	7	US-11-052-554A-284
15	64.5	7.8	463	6	US-10-510-386-186
16	64	7.7	585	6	US-10-821-234-1489
17	63.5	7.7	444	6	US-10-525-710-48
18	63.5	7.7	559	6	US-10-873-528-158
19	63.5	7.7	674	6	US-10-467-657-6812
20	63	7.6	332	6	US-10-793-626-550
21	63	7.6	406	6	US-10-770-726-73
22	62	7.5	514	6	US-10-840-688-2
23	62	7.5	514	6	US-10-840-688-21
24	62	7.5	792	6	US-10-467-657-6026
25	62	7.5	792	6	US-10-467-657-7528

26	62	7.5	887	6	US-10-467-657-3960	Sequence 3960, Ap
27	62	7.5	925	6	US-10-454-437-50	Sequence 50, Appl
28	61.5	7.4	481	7	US-11-090-439-16	Sequence 16, Appl
29	61	7.3	502	6	US-10-821-234-1554	Sequence 1554, Ap
30	61	7.3	514	6	US-10-840-688-3	Sequence 3, Appl
31	60.5	7.3	257	6	US-10-467-657-1522	Sequence 1522, Ap
32	60.5	7.3	950	6	US-10-467-657-854	Sequence 854, App
33	60.5	7.3	2340	7	US-11-052-554A-171	Sequence 171, App
34	60	7.2	514	6	US-10-840-688-10	Sequence 10, Appl
35	60	7.2	514	6	US-10-840-688-12	Sequence 12, Appl
36	60	7.2	523	6	US-10-467-657-2418	Sequence 2418, Ap
37	59.5	7.2	321	7	US-10-959-310-4	Sequence 4, Appl
38	59.5	7.2	321	7	US-11-131-212-72	Sequence 72, Appl
39	59.5	7.2	502	6	US-10-467-657-7332	Sequence 7332, Ap
40	59.5	7.2	1503	7	US-11-087-100-6	Sequence 6, Appl
41	59.5	7.2	1503	7	US-11-087-084-6	Sequence 6, Appl
42	59.5	7.2	1503	7	US-11-087-085-6	Sequence 6, Appl
43	59	7.1	263	7	US-11-082-389-120	Sequence 120, App
44	59	7.1	278	7	US-11-082-389-118	Sequence 118, App
45	59	7.1	338	7	US-11-016-564-11	Sequence 11, Appl

#### ALIGNMENTS

RESULT 1  
US-10-467-657-6110  
; Sequence 6110, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10467.657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 6110  
; LENGTH: 149  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-6110

Query Match	33.3%	Score 276;	DB 6;	Length 149;
Best Local Similarity	41.5%;	Pred. No. 5.1e-23;		
Matches	59;	Conservative	26;	Mismatches 51; Indels 6; Gaps 2;
Qy	17	ILQLPHRYMLLVDRIRIELQANKKIYAYKNITFNEDVFNHFPNKPFPVGLVIEGMAQ	76	
Db	12	IQKLPHRYFLQDRITAFEPKMTLTAIKNVITNEPQFGHFDLPVMPVGLIIEAMAQ	71	
Qy	77	TGGFLATSLWGPDPPEIAKTVIVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGMIAQV	136	
Db	72	ACGTLATLSEGG-----RKNEFFFFAGIDEARFKRQVPGDOLVFEVELLTSRRGIGKFN	127	
Qy	137	GTAQVGGKVVAAELKAMIAER	158	
Db	128	AVAKVGGQVAEVI--MCAKR	147	

RESULT 2  
US-10-467-657-7550  
; Sequence 7550, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita



; PRIOR APPLICATION NUMBER: US 60/141031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: DE 19931636.8  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 199332125.6  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 199332126.4  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 199332127.2  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 199332128.0  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 199332129.9  
; PRIOR FILING DATE: 1999-07-19  
; PRIOR APPLICATION NUMBER: DE 199332226.0  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 199332920.6  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 199332922.2  
; PRIOR FILING DATE: 1999-07-14  
; Remaining prior application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 442  
; SEQ ID NO 324  
; LENGTH: 121  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-10-454-437-324

Query Match 9.1%; Score 75.5; DB 6; Length 121;  
Best Local Similarity 28.7%; Pred. No. 0.17;  
Matches 27; Conservative 15; Mismatches 35; Indels 17; Gaps 5;  
QY 73 GMAQTGGFLAFTSL---WG--FDPEIAKTKIVFMTIDKVKRIPVTPGDRLEY-----H 122  
DB 27 GGAIAHGFLTSLMIPFWGELLDVTGTTKVN--GLDKVFTSPVKVGRIRMGAVRE 84  
QY 123 LEVLKHKGMVQVGTGAQVDGKVAEAEKAMIA 156  
DB 85 ISEVKGNGLHLVADGTITIEGQ-----ERPAAVA 113

RESULT 6  
US-10-485-517-136  
; Sequence 136, Application US/10485517  
; Publication No. US20050256299A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Sheffield  
; APPLICANT: Biosynex Incorporated  
; APPLICANT: Foster, Simon  
; APPLICANT: Mond, James  
; TITLE OF INVENTION: Antigenic Polypeptides  
; FILE REFERENCE: P100629W0  
; CURRENT APPLICATION NUMBER: US/10/485,517  
; CURRENT FILING DATE: 2004-02-02  
; PRIOR APPLICATION NUMBER: GB 0118825.9  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: GB 0200349.9  
; PRIOR FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 424  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 136  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-485-517-136

Query Match 9.0%; Score 74.5; DB 6; Length 485;  
Best Local Similarity 23.8%; Pred. No. 1.3;  
Matches 31; Conservative 21; Mismatches 47; Indels 31; Gaps 6;  
QY 2 ESHONLQSQPFIEHLQILPHRYPMMLVDRIELQ-----ANKKIVAYKNITFNEVDVP 55  
DB 249 BEQVGLNEFFFCG---EILPHRAKALVNRPPSATIVNTYGTETAVTTSIQITQEIL 305

QY 56 NGHPNKP1-----PPGVLIVEGMAQTGGFL-----AFTSLWGFDPPIAKTKIV 99  
DB 306 D-QYPTLVGVVERPGARLSTTDEGBLVIEGQSVSLGYLKNQDKTAEVFNFDGIRT----- 360  
QY 100 YFMTIDKVKF 109  
DB 361 -YHTGDKAKF 369  
RESULT 7  
US-11-194-246-309  
; Sequence 309, Application US/11194246  
; Publication No. US20050272089A1  
; GENERAL INFORMATION:  
; APPLICANT: Mott, John  
; APPLICANT: Trepod, Catherine  
; APPLICANT: Arvidson, Staffan  
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METI  
; TITLE OF INVENTION: USE  
; FILE REFERENCE: 00592.US1 (M&R 268.05920101)  
; CURRENT APPLICATION NUMBER: US/11/194,246  
; CURRENT FILING DATE: 2005-08-01  
; PRIOR APPLICATION NUMBER: US/10/274,586  
; PRIOR FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: US 60/345,438  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 621  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 309  
; LENGTH: 393  
; TYPE: PRT  
; ORGANISM: HAEMOPHILUS INFLUENZAE  
US-11-194-246-309

Query Match 8.9%; Score 73.5; DB 7; Length 393;  
Best Local Similarity 21.6%; Pred. No. 1.3;  
Matches 44; Conservative 24; Mismatches 47; Indels 89; Gaps 10;  
QY 16 HILQILPHRYPMMLVDRIELQANKIVAYKNITFNEVDVFNKGHPFNKPIFFPGVLIVEGM- 74  
DB 189 HIIRDLPFEEY-----KIYSEKNFTFNK--IKQANENGLLMDKTIINVDGMK 231  
QY 75 -----AQTGGFLAFTSL-----WGFI-----D 90  
DB 232 TGHTSQAGYNLVASATTNNMRLISVVMGVPTKYGREVESKKLLQWGFANFETKLEAG 291  
QY 91 PEIAKTKIVY-----FMTIDK-----VKPRIPVTPGDRLEYHLEVLKHKGM 131  
DB 292 KEISEQRVYVYGDKNVVKLGALMDHFITIPKQKQSEVKARYELA-----DKNIQAPLVKQ 346  
QY 132 IWQVGGTA-QVDGKVAEAEKAM 154  
DB 347 V--IGKVYVQLDGKDIASANLQVM 368

RESULT 8  
US-10-821-234-1340  
; Sequence 1340, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0

```
; SEQ ID NO 1340
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1340

Query Match      8.6%; Score 71.5; DB 6; Length 564;
Best Local Similarity 21.7%; Pred. No. 3.4;
Matches 39; Conservative 27; Mismatches 59; Indels 55; Gaps 9;

QY 2 BSHQNLQSQPFIEHLQLPHRYPMMLVDRILIELQANKK-----IV 43
Db 365 EQMHA-----LLAALATWYPM-RIDESIHLQREKYGDKMLRMQKGDPPVYEELF 413

QY 44 AYKNITP-----NEDVFNHGPNNKPIPGVLI-----VEGMAQTGGFLAFTSLMGDPDEI 93
Db 414 SYSCKPFLSPVPVNDVNPVNHKEPFLQQLKVFSDVQQQLSTIRFLKLYTMP-- 471

QY 94 AKTKIVYFMTIDKVKFRIPVTPGDRLEHLEVLKH--KGMIVQVGTAQVDGKVAEAE 151
Db 472 -VAKLAGFLDLTEQFRI-----QLLVFKHKMKNLVW-TSGISALDGEFQSASEV 519

RESULT 9
US-11-042-988-12
; Sequence 12, Application US/11042988
; Publication No. US20050244818A1
; GENERAL INFORMATION:
; APPLICANT: SILICIANO, ROBERT
; APPLICANT: ZHANG, HAILI
; APPLICANT: ZHOU, YAN
; TITLE OF INVENTION: SINGLE CELL ANALYSIS OF HIV REPLICATION CAPACITY AND
; FILE OF INVENTION: DRUG RESISTANCE
; FILE REFERENCE: 62760(71699)
; CURRENT APPLICATION NUMBER: US/11/042,988
; CURRENT FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/540,716
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 12
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-042-988-12

Query Match      8.5%; Score 70.5; DB 7; Length 912;
Best Local Similarity 23.6%; Pred. No. 8.1;
Matches 37; Conservative 26; Mismatches 45; Indels 49; Gaps 8;

QY 17 ILQLPHRYPMMLVDRILIELQANKK-----IVAYKNITFNEDEVFNCHFPNKPFG--- 67
Db 569 IIAQPDQSESELVNIIEQLIKKEKYVLAWPFAHKGIGNEQV-----DKLVSAGIRK 622

QY 68 VLVVEGMAQT-----GGFLAFTSLMGDPDEIAKTKIVYFMTIDKVKFR-----IP 112
Db 623 VFLDGDIDKAQDEHEKHENWRAMASDFNLPPVVAKEIVA---SCDKCQLKGAMHQCVD 679

QY 113 VTPGDRLEHLEVLKHGMIWQVGTAQVDGKVAEAE 149
Db 680 CSPG-----IWQLDCT-HLEBKGKILVA 700

RESULT 10
US-11-093-746A-26
; Sequence 26, Application US/11093746A
; Publication No. US20050266443A1
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo M.
; APPLICANT: Calin, George A.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE AND
; FILE OF INVENTION: COMPOSITIONS AND METHODS FOR MAKING AND USING THE SAME
; FILE REFERENCE: 3589.1015-003
```

```
; CURRENT APPLICATION NUMBER: US/11/093,746A
; CURRENT FILING DATE: 2005-03-30
; PRIOR APPLICATION NUMBER: PCT/US2003/032270
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/417,842
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-093-746A-26

Query Match      8.0%; Score 66; DB 7; Length 205;
Best Local Similarity 26.0%; Pred. No. 3.6;
Matches 13; Conservative 16; Mismatches 13; Indels 8; Gaps 2;

QY 95 KTKIVYFMTIDKVKFRIPVTPGDRLEHLEVLKHGKM---IWQVGTAQV 141
Db 30 KTTILYKLKLGCVTTVPT-----IGFNLTVYKGINFTVWDIGGQEKI 74

RESULT 11
US-11-103-957-92
; Sequence 92, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-11-103-957-92

Query Match      8.0%; Score 66; DB 7; Length 792;
Best Local Similarity 28.4%; Pred. No. 21;
Matches 27; Conservative 10; Mismatches 28; Indels 30; Gaps 5;

QY 79 GFLAFT-----SLWG--FDPEIAKTKIVYFMTIDK---VKFRIPVTPGDRLE 120
Db 469 GSLSTDPYFTADGVSGLGYDIYKAFDPKASTVKQYKTTTAGGGVRMGIPVTEYDRVN 528

QY 121 Y-----HLEV-----LKHGMIWQVGTAQVDG 143
Db 529 FGLAAEHLTVNTYNKAPRYADFIKQYKTDGADG 563

RESULT 12
US-11-052-554A-363
; Sequence 363, Application US/11052554A
; Publication No. US20050288666A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
```

```
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 363
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-363

Query Match
Best Local Similarity 7.9%; Score 65.5; DB 7; Length 129;
Matches 26; Conservative 18; Mismatches 39; Indels 19; Gaps 4;

Qy 33 IIELOANK--KIVAYKNITNEOVFNHGNPKPIPPQVLVEGMAQ---TGGFLAFTSLW 87
Db 41 IVKTDNKTDEKVAFAKGTATMDVLKDNKVKKSGGFITTDGVTQKKAGRY-----W 94

Qy 88 GFDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEHLEVLKHK 129
Db 95 MFD-----VNDKLASKAADKIKVNGDKIEFYLYKVKYK 128

RESULT 13
US-11-055-822-628
; Sequence 628, Application US/11055822
; Publication No. US2005026070A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 628
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-628

Query Match
Best Local Similarity 7.9%; Score 65.5; DB 7; Length 323;
Matches 28; Conservative 18; Mismatches 38; Indels 35; Gaps 6;

; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 363
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-363

Query Match
Best Local Similarity 7.9%; Score 65.5; DB 7; Length 129;
Matches 26; Conservative 18; Mismatches 39; Indels 19; Gaps 4;

Qy 33 IIELOANK--KIVAYKNITNEOVFNHGNPKPIPPQVLVEGMAQ---TGGFLAFTSLW 87
Db 41 IVKTDNKTDEKVAFAKGTATMDVLKDNKVKKSGGFITTDGVTQKKAGRY-----W 94

Qy 88 GFDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEHLEVLKHK 129
Db 95 MFD-----VNDKLASKAADKIKVNGDKIEFYLYKVKYK 128

RESULT 14
US-11-052-554A-284
; Sequence 284, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 284
; LENGTH: 1343
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-284

Query Match
Best Local Similarity 7.8%; Score 65; DB 7; Length 1343;
Matches 33; Conservative 10; Mismatches 37; Indels 26; Gaps 7;

Qy 67 GVLIVEGMAQT---GGFLAFTSLWGFDFPEIAKT-KIYFMTI---DKVKFRIPVTPGDR 118
Db 924 GVQLVK--AQTVIGSGSLTLDLQGDVEADKTLHIAQNGTVAEGDYGFRLLTAPGDG 981

Qy 119 L--BYHLEVL-----KHGMIWQVGGTAQVQDKVVAEAL 151
Db 982 LYNVYGLKALNHHGGQKLTAAEHGG---AYGATADMSAKIGGDL 1024

RESULT 15
US-10-510-386-186
; Sequence 186, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 186
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-186

Query Match
Best Local Similarity 7.8%; Score 64.5; DB 6; Length 463;
Matches 22; Conservative 14; Mismatches 26; Indels 33; Gaps 5;

Qy 64 IFPGVLIV-----EGMAQTGGFL-----AFTSLWGFDFPEIAKTIVYFM 102
```

Db 49 IIGGILVILIGLYAELSSAIPETGGGLIFVYRAFGRKTAFAAANG-----VLFG 98

Qy 103 TIDKVKFRIPVTPGDRLEYHLEVLXHKHGMWQVGG 137

Db 99 YVSVITFEAVALP-TVIDYVLPV-EHQGFMSLSG 131

Search completed: January 10, 2006, 13:38:18  
Job time : 9 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:35:05 ; Search time 77 Seconds  
(without alignments)  
907.289 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MQSHQNLQSFIEHLQI.....QVDGKVVAELKAMIAERD 159

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2443163 seqs, 439378781 residues

Word size : 8

Total number of hits satisfying chosen parameters: 40

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*
- 9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	100.0	159	6	ABU31058 Protein e
2	118	74.2	159	4	AUU36008 Helicobac
3	12	7.5	139	6	ABU46781 Protein e
4	12	7.5	140	3	AAB03793 Fatty aci
5	12	7.5	140	4	AUU38023 Streptoco
6	12	7.5	140	4	AAM01011 CPG 10 pr
7	12	7.5	140	5	ABP28016 Streptoco
8	12	7.5	140	5	ABP28015 Streptoco
9	12	7.5	140	6	ABU00778 S. pneumo
10	12	7.5	140	6	ABU45885 Protein e
11	12	7.5	140	8	ADK47391 Streptoco
12	12	7.5	140	8	ADV89450 Streptoco
13	12	7.5	140	8	ADV80703 Streptoco
14	12	7.5	140	8	ADV82870 Streptoco
15	12	7.5	146	8	ADR94257 Novel S.
16	12	7.5	146	9	AEA58127 Streptoco
17	12	7.5	160	4	AAE02200 Streptoco
18	11	6.9	141	6	ABU23918 Protein e
19	10	6.3	146	6	ABU26345 Protein e
20	9	5.7	102	6	ABU27427 Protein e
21	9	5.7	102	6	ABU27429 Protein e
22	9	5.7	146	9	AEA49209 L. rhamno
23	9	5.7	151	4	AUU34452 E. coli c
24	9	5.7	151	6	ABU15025 Protein e

25	9	5.7	151	6	ABU23117 Protein e
26	9	5.7	151	8	ADN18054 Bacterial
27	9	5.7	157	8	ADN17554 Bacterial
28	9	5.7	162	7	ADF06342 Bacterial
29	9	5.7	172	6	ABU40649 Protein e
30	9	5.7	172	8	ADS42775 Bacterial
31	9	5.7	181	6	ABU50043 Protein e
32	8	5.0	133	3	AAU90363 S. zooepi
33	8	5.0	137	8	ADN20318 Bacterial
34	8	5.0	139	6	ABU30001 Protein e
35	8	5.0	152	6	ABU39542 Protein e
36	8	5.0	154	8	ADS30417 Bacterial
37	8	5.0	161	3	AAG12580 Zea mays
38	8	5.0	437	8	ADX77098 Planc ful
39	8	5.0	478	8	ADN19962 Bacterial
40	8	5.0	910	8	ADS43079 Bacterial

#### ALIGNMENTS

##### RESULT 1

ABU31058

ID ABU31058 standard; protein; 159 AA.

XX AC ABU31058;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #16585.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Helicobacter pylori.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA34928.

XX PT New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 58982; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 159 AA;

Query Match 100.0%; Score 159; DB 6; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-161; Mismatches 0; Gaps 0;  
 Matches 159; Conservative 0; Indels 0;

QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRILQANKKIVAYKNITFNEVDVFNHFP 60  
 DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRILQANKKIVAYKNITFNEVDVFNHFP 60  
 QY 61 NKPFPFGLVVEGMAQTGGFLAFTSLWGFDEPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120  
 DB 61 NKPFPFGLVVEGMAQTGGFLAFTSLWGFDEPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120  
 QY 121 YHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAERD 159  
 DB 121 YHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAERD 159

RESULT 2  
 AAU36008  
 ID AAU36008 standard; protein; 159 AA.

AC AAU36008;  
 DT 14-FEB-2002 (first entry)  
 XX Helicobacter pylori cellular proliferation protein #321.  
 DE Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.  
 XX Helicobacter pylori.

OS WO200170955-A2.  
 XX 27-SEP-2001.  
 XX 21-MAR-2001; 2001WO-US009180.  
 XX 21-MAR-2000; 2000US-0191078P.  
 XX 23-MAY-2000; 2000US-0206848P.  
 XX 26-MAY-2000; 2000US-0207272P.  
 XX 23-OCT-2000; 2000US-0242578P.  
 XX 27-NOV-2000; 2000US-0253625P.  
 XX 22-DEC-2000; 2000US-0257931P.  
 XX 16-FEB-2001; 2001US-0269308P.

(ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KI, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX

DR WPI; 2001-611495/70.  
 XX N-PSDB; AAS53867.  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.  
 XX Example 3; SEQ ID NO 11601; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*  
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 159 AA;

Query Match 74.2%; Score 118; DB 4; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-117; Mismatches 0; Gaps 0;  
 Matches 118; Conservative 0; Indels 0;

QY 41 KIVAYKNITFNEVDVFNHFPNKPFPGLVVEGMAQTGGFLAFTSLWGFDEPEIAKTIVY 100  
 DB 41 KIVAYKNITFNEVDVFNHFPNKPFPGLVVEGMAQTGGFLAFTSLWGFDEPEIAKTIVY 100  
 QY 101 FMTIDKVKFRIPVTPGDRLEHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAER 158  
 DB 101 FMTIDKVKFRIPVTPGDRLEHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAER 158

RESULT 3  
 ABU46781  
 ID ABU46781 standard; protein; 139 AA.

AC ABU46781;  
 XX 19-JUN-2003 (first entry)  
 XX Protein encoded by prokaryotic essential gene #32308.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX Streptococcus pyogenes.

OS WO200277183-A2.  
 XX 03-OCT-2002.  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 XX 06-SEP-2001; 2001US-00948993.  
 XX 25-OCT-2001; 2001US-0342923P.  
 XX 08-FEB-2002; 2002US-00072851.  
 XX 06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KI, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX

DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA50651.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 25; SEQ ID NO 74705; 1766pp; English.  
 PS  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 139 AA;  
 SQ

Query Match 7.5%; Score 12; DB 6; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 0.0003;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
 |||||  
 DB 10 LPHRYPMLLVDR 21

RESULT 4  
 AAB03793  
 ID AAB03793 standard; protein; 140 AA.  
 XX  
 XX AC AAB03793;  
 XX  
 DT 13-OCT-2000 (first entry)  
 XX  
 DE Fatty acid biosynthetic pathway protein FabZ.  
 XX  
 KW Fatty acid biosynthesis; FabZ; beta-hydroxyacyl-ACP; trans-2-acyl-ACP;  
 KW fatty acid elongation; Streptococcus pneumoniae; antibody.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 XX WO200030662-A1.  
 PN  
 XX 02-JUN-2000.  
 PD  
 XX 09-NOV-1999; 99WO-US026435.  
 XX

PR 19-NOV-1998; 98US-00196388.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 FA  
 XX Warren PV, Konstantinidis A, Russell RB;  
 PI  
 XX WPI; 2000-548599/50.  
 DR N-PSDB; AAS59854.  
 DR  
 XX Streptococcus pneumoniae fabZ proteins useful for diagnosing and treating  
 PT microbial infections.  
 PT  
 XX Claim 1; Page 3; 53pp; English.  
 PS  
 XX This sequence represents the fatty acid biosynthetic pathway protein  
 CC FabZ. The invention relates to FabZ nucleotide and protein sequences.  
 CC FabZ catalyses the dehydration of nascent beta-hydroxyacyl-ACP to trans-2  
 CC -acyl-ACP in the process of fatty acid elongation. The invention also  
 CC includes an antibody specific for the FabZ nucleotide sequence, and a  
 CC method for the treatment of an individual in need of enhanced FabZ  
 CC expression. Also included is a method for diagnosing or prognosing a  
 CC disease or a susceptibility to a disease in an individual related to the  
 CC expression of or activity of a FabZ protein. FabZ nucleotide sequences  
 CC and protein sequences may be used in the prevention, treatment and  
 CC diagnosis of diseases associated with fabZ expression and Streptococcal  
 CC infection. FabZ nucleotide and protein sequences may be used to treat  
 CC diseases by rectifying defects in a genome that affect the activity of  
 CC fabZ. FabZ nucleotide sequences can be used to produce FabZ protein, and  
 CC also may be used as probes and primers in diagnostic assays. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC against fabZ and in assays to identify modulators of fabZ expression and  
 CC activity. The anti-fabZ antibodies and fabZ antagonists may also be used  
 CC to down regulate fabZ expression and activity. They may be used to treat  
 CC *S. pneumoniae* infections. The anti-fabZ antibodies may also be used as  
 CC diagnostic agents for detecting the presence of fabZ polypeptides in  
 CC samples  
 XX  
 SQ Sequence 140 AA;  
 Query Match 7.5%; Score 12; DB 3; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
 |||||  
 DB 11 LPHRYPMLLVDR 22

RESULT 5  
 AAU38023  
 ID AAU38023 standard; protein; 140 AA.  
 XX  
 XX AC AAU38023;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Streptococcus pneumoniae cellular proliferation protein #452.  
 XX  
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO200170955-A2.  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX 21-MAR-2001; 2001WO-US0009180.  
 PF  
 XX 21-MAR-2000; 2000US-0191078P.  
 PR 23-MAY-2000; 2000US-0206848P.  
 PR 26-MAY-2000; 2000US-0207727P.  
 PR 23-OCT-2000; 2000US-0242578P.  
 PR



PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;  
 PI Tettelin H;  
 XX  
 DR WPI; 2002-352536/38.  
 DR N-PSDB; ABN68647.  
 XX  
 PT New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for  
 PT detecting a compound that binds to the protein.  
 XX  
 PS Claim 1; Page 3684; 4525pp; English.  
 XX  
 CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC nucleic acids encoding (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins  
 XX  
 SQ Sequence 140 AA;  
 Query Match 7.5%; Score 12; DB 5; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 21 LPHRYPMLLVDR 32  
 Db 11 LPHRYPMLLVDR 22  
 RESULT 8  
 ABP28015  
 ID ABP28015 standard; protein; 140 AA.  
 AC  
 AC ABP28015;  
 XX  
 DT 02-JUL-2002 (first entry).  
 XX  
 DE Streptococcus polypeptide SEQ ID NO 5206.  
 XX  
 DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW Group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX  
 OS Streptococcus agalactiae.  
 XX  
 PN WO200234771-A2.  
 XX  
 PD 02-MAY-2002.  
 XX  
 PF 29-OCT-2001; 2001WO-GB004789.  
 XX  
 PF 27-OCT-2000; 2000GB-00026333.  
 PR 24-NOV-2000; 2000GB-00028727.  
 PR 07-MAR-2001; 2001GB-00005640.  
 XX  
 PA (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;  
 PI Tettelin H;  
 XX

DR WPI; 2002-352536/38.  
 DR N-PSDB; ABN68646.  
 XX  
 PT New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for  
 PT detecting a compound that binds to the protein.  
 XX  
 PS Claim 1; Page 3684; 4525pp; English.  
 XX  
 CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC nucleic acids encoding (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins  
 XX  
 SQ Sequence 140 AA;  
 Query Match 7.5%; Score 12; DB 5; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 21 LPHRYPMLLVDR 32  
 Db 11 LPHRYPMLLVDR 22  
 RESULT 9  
 ABU00778  
 ID ABU00778 standard; protein; 140 AA.  
 AC  
 AC ABU00778;  
 XX  
 XX 23-OCT-2003 (revised)  
 DT 11-FEB-2003 (first entry)  
 XX  
 DE S. pneumoniae type 4 strain protein from coding region #346.  
 XX  
 DE Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
 KW gene therapy; vaccine.  
 XX  
 OS Streptococcus pneumoniae; type 4 strain.  
 XX  
 PN WO200277021-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-IB002163.  
 XX  
 PF 27-MAR-2001; 2001GB-00007658.  
 XX  
 PA (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Massignani V, Tettelin H, Fraser C;  
 XX  
 DR WPI; 2003-040579/03.  
 DR N-PSDB; ABX06056.  
 XX  
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
 PT useful as medicaments for treating or preventing a disease or infection



PD 02-MAR-2004.  
XX 26-MAY-2000; 2000US-00583110.  
XX  
XX 02-JUL-1997; 97US-0051553P.  
PR 12-MAY-1998; 98US-0085131P.  
PR 30-JUN-1998; 98US-00107433.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CB;  
PI WPI; 2004-212399/20.  
DR N-PSDB; ADR44730.  
XX  
XX New nucleic acid molecules and polypeptides useful for diagnosing,  
PT preventing and treating pathological conditions resulting from bacterial  
PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
PT screening.  
XX  
XX Disclosure; SEQ ID NO 3906; 301pp; English.  
XX  
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
CC and polypeptides. The nucleic acids and proteins are useful for  
CC diagnosing, preventing and treating pathological conditions resulting  
CC from bacterial infection, such as S. pneumoniae infection. These may also  
CC be used for drug screening procedures. The present sequence represents a  
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
CC data for this patent did not appear in the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html.  
XX  
XX  
SQ Sequence 140 AA;  
  
Query Match 7.5%; Score 12; DB 8; Length 140;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 21 LPHRYPMLLVDR 32  
DB 11 LPHRYPMLLVDR 22  
|||||  
  
RESULT 12  
ADV89450  
ID ADV89450 standard; protein; 140 AA.  
XX  
XX ADV89450;  
AC  
XX  
XX 24-FEB-2005 (first entry)  
DT  
XX  
DE Streptococcus agalactiae protein sequence, SEQ ID 1844.  
XX  
XX Antibacterial; Vaccine; bacterial infection.  
KW  
XX  
XX Streptococcus agalactiae.  
OS  
XX  
XX WO200292818-A2.  
PN  
XX  
XX 21-NOV-2002.  
PD  
XX  
XX 26-APR-2002; 2002WO-IB003059.  
PF  
XX  
XX 26-APR-2001; 2001FR-00005642.  
PR  
XX  
XX (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;  
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Quot P, Kunst F;  
XX WPI; 2004-101891/11.  
DR  
XX  
XX Genomic nucleotide sequences encoding polypeptides of Streptococcus  
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips  
PT and identification of therapeutic targets.  
XX  
XX Claim 6; SEQ ID NO 1844; 439pp; French.  
XX  
XX The present invention relates to novel Streptococcus agalactiae  
CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476), and  
CC nucleotide sequences (II; ADV7746-ADV89950). The nucleotide sequences encode polypeptides of S.  
agalactiae involved in the synthesis of amino acids, cell membranes,  
intermediate (central) metabolism, energetic metabolism, fatty acid and  
phospholipid metabolism, nucleotide metabolism including purines,  
pyrimidines and/or nucleosides, regulatory functions, replication,  
transcription, translation, protein transport, adaptation to atypical  
conditions, sensitivity to medicines and/or analogues, functions related  
to transporters, biosynthesis of cofactors, prosthetic groups and  
cellular machinery. (I) are useful for the detection and/or amplification of nucleic acids.  
Pharmaceutical composition comprising (I) or (II) are useful for  
treatment of a bacterial S. agalactiae infection. Note: WO200292818A2 is  
equivalent for the present basic patent FR2824074A1. WO200292818A2  
contains 6617 sequence whereas the present patent only contains 2344  
sequences.

PT Genomic nucleotide sequences encoding polypeptides of Streptococcus  
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips  
XX and identification of therapeutic targets.  
XX  
XX Claim 6; SEQ ID NO 1844; 2687pp; French.  
XX  
XX The present invention relates to novel Streptococcus agalactiae  
CC nucleotide sequences (I; ADV87607-ADV87745) and novel polypeptides (II;  
CC ADV87746-ADV89950). The nucleotide sequences encode polypeptides of S.  
CC agalactiae involved in the synthesis of amino acids, cell membranes,  
CC intermediate (central) metabolism, energetic metabolism, fatty acid and  
CC phospholipid metabolism, nucleotide metabolism including purines,  
CC pyrimidines and/or nucleosides, regulatory functions, replication,  
CC transcription, translation, protein transport, adaptation to atypical  
CC conditions, sensitivity to medicines and/or analogues, functions related  
CC to transporters, biosynthesis of cofactors, prosthetic groups and  
CC cellular machinery. (I) are useful for the detection and/or amplification of nucleic acids.  
CC Pharmaceutical composition comprising (I) or (II) are useful for  
CC treatment of a bacterial S. agalactiae infection. Note: WO200292818A2 is  
CC equivalent for the present basic patent FR2824074A1. WO200292818A2  
CC contains 6617 sequence whereas the present patent only contains 2344  
CC sequences.  
XX  
XX Sequence 140 AA;  
  
Query Match 7.5%; Score 12; DB 8; Length 140;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 21 LPHRYPMLLVDR 32  
DB 11 LPHRYPMLLVDR 22  
|||||  
  
RESULT 13  
ADV80703  
ID ADV80703 standard; protein; 140 AA.  
XX  
XX ADV80703;  
AC  
XX  
XX 24-FEB-2005 (first entry)  
DT  
XX  
DE Streptococcus agalactiae protein, SEQ ID 1844.  
XX  
XX Antibacterial; vaccine; bacterial infection.  
KW  
XX  
XX Streptococcus agalactiae.  
OS  
XX  
XX WO200292818-A2.  
PN  
XX  
XX 21-NOV-2002.  
PD  
XX  
XX 26-APR-2002; 2002WO-IB003059.  
PF  
XX  
XX 26-APR-2001; 2001FR-00005642.  
PR  
XX  
XX (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;  
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Quot P, Kunst F;  
XX WPI; 2004-101891/11.  
DR  
XX  
XX Genomic nucleotide sequences encoding polypeptides of Streptococcus  
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips  
PT and identification of therapeutic targets.  
XX  
XX Claim 6; SEQ ID NO 1844; 439pp; French.  
XX  
XX The present invention relates to novel Streptococcus agalactiae  
CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476), and  
CC nucleotide sequences (II; ADV7746-ADV89950). The nucleotide sequences encode polypeptides of S.  
agalactiae involved in the synthesis of amino acids, cell membranes,  
intermediate (central) metabolism, energetic metabolism, fatty acid and  
phospholipid metabolism, nucleotide metabolism including purines,  
pyrimidines and/or nucleosides, regulatory functions, replication,  
transcription, translation, protein transport, adaptation to atypical  
conditions, sensitivity to medicines and/or analogues, functions related  
to transporters, biosynthesis of cofactors, prosthetic groups and  
cellular machinery. (I) are useful for the detection and/or amplification of nucleic acids.  
Pharmaceutical composition comprising (I) or (II) are useful for  
treatment of a bacterial S. agalactiae infection. Note: WO200292818A2 is  
equivalent for the present basic patent FR2824074A1. WO200292818A2  
contains 6617 sequence whereas the present patent only contains 2344  
sequences.



CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The  
 CC nucleotide sequences encode polypeptides of *S. agalactiae* involved in the  
 CC synthesis of amino acids, cell membranes, intermediate (central)  
 CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,  
 CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,  
 CC regulatory functions, replication, transcription, translation, protein  
 CC transport, adaptation to atypical conditions, sensitivity to medicines  
 CC and/or analogues, functions related to transposons, biosynthesis of  
 CC cofactors, prosthetic groups and transporters, cell membrane proteins and  
 CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are  
 CC useful for treatment of a bacterial *S. agalactiae* infection. The complete  
 CC genome of *Streptococcus agalactiae* is given in ADV81204. Note: The  
 CC present patent is an equivalent for the basic patent FR2824074A1, which  
 CC contains only 2344 sequences.

XX Sequence 140 AA;

Query Match 7.5%; Score 12; DB 8; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
 |||||  
 DB 11 LPHRYPMLLVDR 22

#### RESULT 14

ADV82870  
 ID ADV82870 standard; protein; 140 AA.

AC ADV82870;

DT 24-FEB-2005 (first entry)

DE Streptococcus agalactiae protein, SEQ ID 4011.

KW Antibacterial; vaccine; bacterial infection.

OS Streptococcus agalactiae.

XX WO200292818-A2.

PN 21-NOV-2002.

PD 26-APR-2002; 2002WO-IB003059.

XX 26-APR-2001; 2001FR-00005642.

PR (INSP ) INST PASTEUR.

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;  
 PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst P;  
 XX WPI; 2004-101891/11.

DR Claim 6; SEQ ID NO 4011; 439pp; French.

XX Genomic nucleotide sequences encoding polypeptides of *Streptococcus*  
 PT agalactiae for the development of vaccines, diagnostic tools, DNA chips  
 PT and identification of therapeutic targets.

XX

XX The present invention relates to novel *Streptococcus agalactiae*  
 CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and  
 CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The  
 CC nucleotide sequences encode polypeptides of *S. agalactiae* involved in the  
 CC synthesis of amino acids, cell membranes, intermediate (central)  
 CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,  
 CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,  
 CC regulatory functions, replication, transcription, translation, protein  
 CC transport, adaptation to atypical conditions, sensitivity to medicines  
 CC and/or analogues, functions related to transposons, biosynthesis of

CC cofactors, prosthetic groups and transporters, cell membrane proteins and  
 CC cellular machinery. (I) are useful for the detection and/or amplification  
 CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are  
 CC useful for treatment of a bacterial *S. agalactiae* infection. The complete  
 CC genome of *Streptococcus agalactiae* is given in ADV81204. Note: The  
 CC present patent is an equivalent for the basic patent FR2824074A1, which  
 CC contains only 2344 sequences.

XX Sequence 140 AA;

Query Match 7.5%; Score 12; DB 8; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
 |||||  
 DB 11 LPHRYPMLLVDR 22

#### RESULT 15

ADR94257  
 ID ADR94257 standard; protein; 146 AA.

XX ADR94257;

XX 16-DEC-2004 (first entry)

DE Novel *S. pneumoniae* protein sequence, SEQ ID 2892.

XX Meningitis; bacteraemia; pneumonia; otitis media; vaccine;  
 KW bacterial infection.

XX Streptococcus pneumoniae.

XX US6800744-B1.

XX 05-OCT-2004.

XX 30-JUN-1998; 98US-00107433.

XX 02-JUL-1997; 97US-0051553P.

PR 12-MAY-1998; 98US-0085131P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2004-697205/68.

DR N-PSDB; ADR91654.

XX New isolated nucleic acid encoding a *Streptococcus pneumoniae*  
 PT polypeptide, useful for diagnosing, preventing and/or treating  
 PT pathological conditions resulting from the bacterial infection.

XX Disclosure; SEQ ID NO 2892; 151pp; English.

XX The invention relates to an isolated nucleic acid comprising a sequence  
 CC encoding a *Streptococcus pneumoniae* ADR91366polypeptide, or its  
 CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,  
 CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,  
 CC ADR96079) or any of the fully defined sequences appearing as ADR91705,  
 CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or  
 CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide  
 CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is  
 CC hybridisable under high stringency conditions to the nucleotide sequence.  
 CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.  
 CC Also included are a recombinant expression vector comprising the isolated  
 CC nucleic acid cited above operably linked to a transcription regulatory  
 CC element, a cell comprising the recombinant expression vector and a probe  
 CC comprising at least 20 consecutive nucleotides of the nucleotide  
 CC sequences as cited above. The methods and compositions of the present  
 CC invention are useful for the diagnosis, prevention and/or treatment of  
 CC pathological conditions resulting from bacterial infection by

CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and  
 CC otitis media. The present sequence is one of the 2603 disclosed S.  
 CC pneumoniae protein sequences. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.  
 XX  
 SQ Sequence 146 AA;

Query Match 7.5%; Score 12; DB 8; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 0.00032;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMMLVDR 32  
 DB 17 LPHRYPMMLVDR 28  
 |||||  
 |||||

RESULT 16  
 AEA58127  
 ID AEA58127 standard; protein; 146 AA.  
 XX AC AEA58127;  
 XX DT 25-AUG-2005 (first entry)  
 XX DE Streptococcus pneumoniae ORF amino acid sequence SEQ ID NO:2892.  
 XX KW bacterial infection; Streptococcus pneumoniae infection; antibacterial;  
 XX KW vaccine.  
 XX OS Streptococcus pneumoniae.  
 XX PN US2005136404-A1.  
 XX PD 23-JUN-2005.  
 XX PF 10-JUL-2003; 2003US-00617320.  
 XX PR 02-JUL-1997; 97US-0051553P.  
 XX PR 12-MAY-1998; 98US-0085131P.  
 XX PR 30-JUN-1998; 98US-00107433.  
 XX PA (DOUC/) DOUCETTE-STAMM L A.  
 XX PA (BUSH/) BUSH D.  
 XX PI Doucette-Stamm LA, Bush D;  
 XX PS WPI; 2005-477576/48.  
 XX DR N-PSDB; AEA55524.  
 XX PT New isolated nucleic acid molecules and encoded polypeptides useful for  
 PT diagnosing, preventing or treating bacterial infections, particularly  
 PT Streptococcus pneumoniae infection.  
 XX Claim 5; SEQ ID NO 2892; 144pp; English.  
 XX The invention relates to an isolated nucleic acid molecule for detecting,  
 CC preventing or treating pathological conditions resulting from bacterial  
 CC infection. The isolated nucleic acid comprises: (a) any of the 2603  
 CC nucleotide sequences of AEA55236 to AEA57838; (b) a nucleotide sequence  
 CC encoding a Streptococcus pneumoniae polypeptide comprising any of the  
 CC 2603 amino acid sequences of AEA57839 to AEA60441; or (c) a nucleotide  
 CC sequence of at least 8 nucleotides in length, where the sequence is  
 CC hybridizable to a nucleic acid having any of the nucleotide sequences in  
 CC (a). Also described: (1) a recombinant expression vector comprising the  
 CC above nucleic acid operably linked to a transcription regulatory element;  
 CC (2) a cell comprising the recombinant expression vector; (3) producing an  
 CC S. pneumoniae polypeptide; (4) a probe comprising a nucleotide sequence  
 CC consisting of at least 8 nucleotides of any of AEA55236 to AEA57838; (5)  
 CC treating a subject for S. pneumoniae infection; (6) a recombinant or  
 CC substantially pure preparation of an S. pneumoniae polypeptide or its  
 CC fragment, where the polypeptide is selected from AEA57839 to AEA60441;

CC (7) a vaccine composition for preventing or treating an S. pneumoniae  
 CC infection, comprising an amount of the above nucleic acid or polypeptide;  
 CC (8) detecting the presence of a Streptococcus nucleic acid in a sample;  
 CC (9) a computer readable medium having recorded the nucleotide sequences  
 CC of AEA55236 to AEA57838; (10) a computer based system for identifying  
 CC fragments of the Streptococcus genome of commercial importance. The  
 CC composition and methods are useful for diagnosing, preventing or treating  
 CC bacterial infections, particularly S. pneumoniae infection. The present  
 CC sequence represents a S. pneumoniae ORF amino acid sequence from the  
 CC present invention. Note - The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from the USPTO web site.  
 XX  
 SQ Sequence 146 AA;

Query Match 7.5%; Score 12; DB 9; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 0.00032;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMMLVDR 32  
 DB 17 LPHRYPMMLVDR 28  
 |||||  
 |||||

RESULT 17  
 AAE02200  
 ID AAE02200 standard; protein; 160 AA.  
 XX AC AAE02200;  
 XX DT 11-SEP-2003 (revised)  
 XX DT 31-JUL-2001 (first entry)  
 XX DE Streptococcus pneumoniae His6-beta-hydroxyacyl-ACP dehydrase (FabZ).  
 XX KW His6-FabZ; high throughput method; fatty acid biosynthesis; therapy;  
 KW bacterial enzyme; biological agent screening; otitis media; empyema;  
 KW bacterial tracheitis; acute epiglottitis; thyroiditis; lung abscess;  
 KW infective endocarditis; secretory diarrhoea; splenic; retroperitoneal;  
 KW intrarenal; perinephric; cerebral; cutaneous; abscess; blepharitis;  
 KW conjunctivitis; keratitis; endophthalmitis; preseptal; orbital;  
 KW cellulitis; dacryocystitis; epididymitis; abscess; toxic shock syndrome;  
 KW impetigo; folliculitis; wound infection; bacterial myositis;  
 KW septic arthritis; osteomyelitis; beta-hydroxyacyl-ACP dehydrase;  
 KW acyl carrier protein.  
 XX OS Streptococcus pneumoniae; 0100993.  
 XX PN WO200130988-A1.  
 XX PD 03-MAY-2001.  
 XX PF 26-OCT-2000; 2000WO-US029451.  
 XX PR 27-OCT-1999; 99US-0161775P.  
 XX PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX PI Dewolf W, Kallender H, Lonsdale JT;  
 XX WPI; 2001-316332/33.  
 XX DR N-PSDB; AAD06211.  
 XX High throughput method for screening for biological agents against fatty  
 PT acid biosynthesis comprises contacting a bacterial enzymatic pathway with  
 PT enzymes e.g. malonyl-CoA:ACP transacylase.  
 XX Claim 1; Page 13; 94pp; English.  
 XX The present invention relates to a high throughput method for screening  
 CC biological agents affecting fatty acid biosynthesis, comprises contacting  
 CC a bacterial enzymatic pathway with enzymes. The method is used for

CC screening biological agents affecting fatty acid biosynthesis. Agonists  
 CC and antagonists of fab (fatty acid biosynthesis) are used to inhibit,  
 CC prevent or treat diseases such as infections of the upper respiratory  
 CC tract (e.g. otitis media, bacterial tracheitis, acute epiglottitis,  
 CC thyroiditis), lower respiratory (e.g. empyema, lung abscess), cardiac  
 CC (e.g. infective endocarditis), gastrointestinal (e.g. secretory  
 CC diarrhoea, splenic abscess, retroperitoneal abscess), central nervous  
 CC system (e.g. cerebral abscess), eye (e.g. blepharitis, conjunctivitis,  
 CC keratitis, endophthalmitis, preseptal and orbital cellulitis,  
 CC dacryocystitis), kidney and urinary tract (e.g. epididymitis, intrarenal  
 CC and perinephric abscess, toxic shock syndrome), skin (e.g. impetigo,  
 CC folliculitis, cutaneous abscess, cellulitis, wound infection, bacterial  
 CC myositis) and bone and joint (e.g. septic arthritis, osteomyelitis). The  
 CC present sequence is Streptococcus pneumoniae 010093 His6-beta-  
 CC hydroxyl-ACP (acyl carrier protein) dehydrase (Fab2). In fatty acid  
 CC biosynthetic pathway, dehydration by Fab2 leads to trans-2-enoyl-ACP  
 CC which is in turn converted to acyl-ACP by enoyl-ACP reductase (FabI).  
 CC (Updated on 11-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 160 AA;

Query Match 7.5%; Score 12; DB 4; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 0.00035; Mismatches 0; Indels 0; Gaps 0;  
 Matches 12; Conservative 0;

QY 21 LPHRYPMLLVDR 32  
 |||||  
 DB 31 LPHRYPMLLVDR 42

RESULT 18  
 ABU23918  
 ID ABU23918 standard; protein; 141 AA.  
 XX  
 AC ABU23918;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #9445.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Clostridium acetobutylicum.  
 XX  
 EN WO20027183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA27788.  
 XX

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 51842; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC on a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 141 AA;

Query Match 6.9%; Score 11; DB 6; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 0.0036;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PHRYPMLLVDR 32  
 |||||  
 DB 13 PHRYPMLLVDR 23

RESULT 19  
 ABU26345  
 ID ABU26345 standard; protein; 146 AA.  
 XX  
 AC ABU26345;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #11872.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Campylobacter jejuni.  
 XX WO20027183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA30215.

XX PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX  
 XX Claim 25; SEQ ID NO 54269; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 146 AA;

Query Match 6.3%; Score 10; DB 6; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 0.044;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAAEAEKAMI 155  
 DB 134 VAAEAEKAMI 143

## RESULT 20

ABU27427  
 ID ABU27427 standard; protein; 102 AA.

XX AC ABU27427;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #12954.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Enterobacter cloacae.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PD 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDS; ACA31297.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 25; SEQ ID NO 55351; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 102 AA;

Query Match 5.7%; Score 9; DB 6; Length 102;

Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70

DB 74 KPFPFGVLI 82

## RESULT 21

ABU27429

ID ABU27429 standard; protein; 102 AA.

XX AC ABU27429;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #12956.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Enterobacter cloacae.

XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR N-PSDB; ACA31299.  
XX DR WPI; 2003-029926/02.  
XX DR N-PSDB; ACA31299.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 55353; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: The sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained  
XX CC in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 102 AA;  
Query Match 5.7%; Score 9; DB 6; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.38; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 62 KPFFPGVLI 70  
Db 74 KPFFPGVLI 82  
RESULT 22  
AEA49209  
ID AEA49209 standard; protein; 146 AA.

XX AC AEA49209;  
XX DT 25-AUG-2005 (first entry)  
XX DE L. rhamnosus polypeptide #73.  
XX KW Gene fusion; wound healing; immunostimulant; antilipemic; vulnerary.  
XX OS Lactobacillus rhamnosus.  
XX PN WO2005056801-A1.  
XX PD 23-JUN-2005.  
XX PF 15-DEC-2003; 2003WO-NZ000278.  
XX PR 15-DEC-2003; 2003WO-NZ000278.  
XX PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX PA (FONT-) FONTEREA COOP GROUP LTD.  
XX PI Glenn M, Havukkala I, Lubbers MW, Dekker J;  
XX DR WPI; 2005-435589/44.  
XX DR N-PSDB; AEA49117.  
XX PS Claim 18; SEQ ID NO 153; 323pp; English.  
XX CC The invention relates to an isolated Lactobacillus rhamnosus  
XX CC polynucleotide and the polypeptide it encodes. The invention also relates  
XX CC to a genetic construct comprising the polynucleotide, a transgenic host  
XX CC cell comprising the genetic construct, a transgenic organism comprising  
XX CC the host cell or its progeny and a method of modulating the activity of a  
XX CC polypeptide in an organism comprising stably incorporating the  
XX CC polynucleotide into the genome of the organism. The polynucleotide is  
XX CC useful for modulating the activity of a polypeptide in an organism, for  
XX CC creating a transgenic organism and for the detection of lactic acid  
XX CC bacteria in a sample. The polynucleotides and polypeptides may be used in  
XX CC the selection and production of more effective probiotic bacteria, for  
XX CC immune function enhancement, for reduction of blood lipids such as  
XX CC cholesterol, for wound healing and in vaccine development. They are also  
XX CC useful as health supplements, as animal probiotics for improved animal  
XX CC health and for the production of improved food bacteria through genetic  
XX CC modification. The polynucleotide may also be useful in genome mapping and  
XX CC physical mapping. This sequence represents an *L. rhamnosus* polypeptide of  
XX CC the invention.  
XX SQ Sequence 146 AA;  
Query Match 5.7%; Score 9; DB 9; Length 146;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 20 ILPHRYPML 28  
Db 16 ILPHRYPML 24  
RESULT 23  
AAU34452  
ID AAU34452 standard; protein; 151 AA.  
XX AC AAU34452;  
XX DT 14-FEB-2002 (first entry)  
XX DE E. coli cellular proliferation protein #33.  
XX DE

KW Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.  
 XX Escherichia coli.  
 OS WO200170955-A2.  
 PN 27-SEP-2001.  
 PD  
 XX  
 XX 21-MAR-2001; 2001WO-US009180.  
 PF  
 XX 21-MAR-2000; 2000US-0191078P.  
 PR 23-MAY-2000; 2000US-0206848P.  
 PR 26-MAY-2000; 2000US-0207272P.  
 PR 23-OCT-2000; 2000US-0242578P.  
 PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 PI  
 XX WPI: 2001-611495/70.  
 DR N-PSDB; AAS52311.  
 DR  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.  
 PT  
 XX Example 3; SEQ ID NO 10045; 511pp; English.  
 PS  
 XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes,  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 151 AA;

Query Match 5.7%; Score 9; DB 4; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 0.53;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFFPGVLI 70  
 DB 58 KPFFPGVLI 66

## RESULT 24

ABU15025  
 ID ABU15025 standard; protein; 151 AA.  
 AC  
 XX ABU15025;  
 XX 19-JUN-2003 (first entry)  
 DT  
 XX Protein encoded by Prokaryotic essential gene #552.  
 DE  
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 KW

XX Escherichia coli.  
 OS WO200277183-A2.  
 PN 03-OCT-2002.  
 PD  
 XX 21-MAR-2002; 2002WO-US009107.  
 PF  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 PI  
 XX WPI: 2003-029926/02.  
 DR N-PSDB; ACA18895.  
 DR  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PT  
 XX Claim 25; SEQ ID NO 42949; 1766pp; English.  
 PS  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 5213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than S. aureus, S. typhimurium,  
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 151 AA;

Query Match 5.7%; Score 9; DB 6; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 0.53;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFFPGVLI 70  
 DB 58 KPFFPGVLI 66

## RESULT 25





CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 151 AA;

Query Match 5.7%; Score 9; DB 8; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 0.53;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPQGLI 70  
 |||||  
 Db 58 KPIFPQGLI 66

RESULT 27  
 ADN17554  
 ID ADN17554 standard; protein; 157 AA.  
 XX  
 AC ADN17554;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX Bacterial polypeptide #207.  
 DE  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX Bacteria.  
 OS  
 XX US2003233675-A1.  
 PN  
 XX 18-DEC-2003.  
 PD  
 XX 20-FEB-2003; 2003US-00369493.  
 PF  
 XX 21-FEB-2002; 2002US-0360039P.  
 PR  
 XX (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 PA (GOLD/) GOLDMAN B S.  
 PI  
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 PI WPI; 2004-061375/06.  
 DR  
 XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.  
 XX  
 PS Claim 1; SEQ ID NO 207; 122pp; English.

The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 157 AA;

Query Match 5.7%; Score 9; DB 8; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 0.55;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPQGLI 70  
 |||||  
 Db 64 KPIFPQGLI 72

RESULT 28  
 ADF06342  
 ID ADF06342 standard; protein; 162 AA.  
 XX  
 AC ADF06342;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX Bacterial polypeptide #2455.  
 DE  
 KW Proteus mirabilis infection; bacterial infection; antibacterial;  
 KW immunostimulant.  
 XX  
 OS Proteus mirabilis.  
 XX  
 PN US6605709-B1.  
 XX  
 PD 12-AUG-2003.  
 XX  
 PF 05-APR-2000; 2000US-00543681.  
 XX  
 PR 09-APR-1999; 99US-0128706P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton GL;  
 XX  
 DR WPI; 2003-895291/82.  
 DR N-PSDB; ADF02170.  
 XX  
 PT New Proteus mirabilis polypeptides and polynucleotides, useful as  
 PT reagents for diagnosis of bacterial disease, as components of  
 PT antibacterial vaccines, as targets for antibacterial drugs, or as  
 PT biocontrol agents for plants.  
 XX  
 PS Disclosure; SEQ ID NO 6627; 870pp; English.

The invention relates to new Proteus mirabilis polypeptides and  
 CC polynucleotides. The invention also relates to antibodies against the  
 CC polypeptides, methods for producing the polypeptides, a method of  
 CC generating vaccines for immunising an individual against P. mirabilis, a  
 CC method for evaluating a compound for the ability to bind a P. mirabilis  
 CC polypeptide and a method for screening test compounds for anti-bacterial  
 CC activity. The polypeptides and polynucleotides are useful as molecular  
 CC targets for diagnosing, preventing and treating pathological conditions  
 CC resulting from bacterial infection, as reagents for diagnosis of  
 CC bacterial diseases, as components of antibacterial vaccines, as targets  
 CC for antibacterial drugs or as bio-control agents for plants. This  
 CC sequence represents a Proteus mirabilis polypeptide of the invention.



SQ Sequence 162 AA;

Query Match 5.7%; Score 9; DB 7; Length 162;  
Best Local Similarity 100.0%; Pred. No. 0.57;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70  
Db 69 KPIPPGVLI 77  
|||||

RESULT 29  
ABU40649  
ID ABU40649 standard; protein; 172 AA.  
XX  
AC ABU40649;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #26176.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Proteus sp.  
XX  
FN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
(ELIT-) ELITRA PHARM INC.  
PA  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
WPI; 2003-029926/02.  
DR N-PSDB; ACA44519.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
Claim 25; SEQ ID NO 68573; 1766pp; English.  
XX  
The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC on a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 172 AA;

Query Match 5.7%; Score 9; DB 6; Length 172;  
Best Local Similarity 100.0%; Pred. No. 0.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70  
Db 81 KPIPPGVLI 89  
|||||

RESULT 30  
ADS42775  
ID ADS42775 standard; protein; 172 AA.  
XX  
AC ADS42775;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #21205.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
FN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PT 21-FEB-2002; 2002US-0360039P.  
XX  
(CAOV/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
DR  
XX New recombinant DNA construct comprising a promoter positioned to provide  
XX for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
PT  
XX Claim 1; SEQ ID NO 21205; 122pp; English.  
PS  
XX The invention relates to a recombinant DNA construct comprising a  
XX promoter functional in a plant cell, where the promoter is positioned to  
XX provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC tolerance to herbicides, e.g. improved cold, heat or drought tolerance,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 172 AA;  
  
Query Match 5.7%; Score 9; DB 8; Length 172;  
Best Local Similarity 100.0%; Pred. No. 0.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 62 KPFFPGVLI 70  
Db 79 KPFFPGVLI 87  
|||||  
|||||  
  
RESULT 31  
ABUS0043  
ID ABUS0043 standard; protein; 181 AA.  
XX  
AC ABUS0043;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #35570.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
OS Yersinia pestis.  
XX  
FN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA53913.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 77967; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway; (8)  
CC required for proliferation, or that inhibits cellular proliferation; (9)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 181 AA;  
  
Query Match 5.7%; Score 9; DB 6; Length 181;  
Best Local Similarity 100.0%; Pred. No. 0.63;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 62 KPFFPGVLI 70  
Db 58 KPFFPGVLI 66  
|||||  
|||||  
  
RESULT 32  
AAY90363  
ID AAY90363 standard; protein; 133 AA.  
XX  
AC AAY90363;  
XX  
DT 12-SEP-2003 (revised)  
DT 15-JAN-2001 (first entry)  
XX  
DE *S. zoepidemicus* gene clone #2 protein #1.  
XX  
KW DNA cloning; gene cloning; gene detection.  
XX  
OS *Streptococcus equi* subsp. *zoepidemicus*.  
XX  
PN EP1035202-A2.  
XX  
PD 13-SEP-2000.  
XX  
PF 09-MAR-2000; 2000EP-00104485.  
XX  
PR 11-MAR-1999; 99JP-00064432.  
XX  
PA (JCRP-) JCR PHARM CO LTD.  
XX  
PI Hong K;  
XX  
DR WPI; 2000-559872/52.  
DR N-PSDB; AAA37788.  
XX  
PT Cloning unknown gene fragments by introducing cloning vector with ligated  
PT DNA PCR products containing restriction enzyme cleaved and  
PT dephosphorylated plasmid DNA and chromosomal DNA into competent cells.  
XX  
PS Claim 5; Page 10-12; 14pp; English.



PS Claim 25; SEQ ID NO 57925; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC on a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 139 AA;  
Query Match 5.0%; Score 8; DB 6; Length 139;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 22 PHRYPMLL 29  
Db 11 PHRYPMLL 18  
|||||  
|||||  
RESULT 35  
ABU39542  
ID ABU39542 standard; protein; 152 AA.  
XX  
XX AC ABU39542;  
XX  
XX DT 19-JUN-2003 (first entry)  
XX  
XX DE Protein encoded by Prokaryotic essential gene #25069.  
XX  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
XX OS Pasteurella multocida.  
XX  
XX PN WO200277183-A2.  
XX  
XX PD 03-OCT-2002.  
XX  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX  
XX PR 21-MAR-2001; 2001US-00815242.  
XX  
XX PR 06-SEP-2001; 2001US-00948993.  
XX  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX  
XX PR 08-FEB-2002; 2002US-00072851.  
XX  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA43412.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 67466; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC on a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 152 AA;  
Query Match 5.0%; Score 8; DB 6; Length 152;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 63 PIFPGVLI 70  
Db 62 PIFPGVLI 69  
|||||  
|||||  
RESULT 36  
ADS30417  
ID ADS30417 standard; protein; 154 AA.  
XX  
XX AC ADS30417;  
XX  
XX DT 02-DEC-2004 (first entry)  
XX  
XX DE Bacterial polypeptide #19450.  
XX  
XX KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.

XX OS Bacteria.  
XX PN US2003233675-A1.  
XX XX 18-DEC-2003.  
XX PF 20-FEB-2003; 2003US-00369493.  
XX PD 21-FEB-2002; 2002US-0360039P.  
XX PA (CAOY/) CAO Y.  
XX PA (HINK/) HINKLE G J.  
XX PA (SLAT/) SLATER S C.  
XX PA (CHEN/) CHEN X.  
XX PA (GOLD/) GOLDMAN B S.  
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX DR WPI; 2004-061375/06.  
XX PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX PS Claim 1; SEQ ID NO 19450; 122pp; English.  
XX CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX SQ Sequence 154 AA;  
  
Query Match 5.0%; Score 8; DB 8; Length 154;  
Best Local Similarity 100.0%; Pred. No. 6.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 21 LPHRYPML 28  
| | | | |  
Db 17 LPHRYPML 24  
  
RESULT 37  
AAG12580  
ID AAG12580 standard; protein; 161 AA.  
XX AC AAG12580;  
XX DT 17-OCT-2000 (first entry)  
XX DE Zea mays protein fragment SEQ ID NO: 11749.  
XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence; corn.  
OS Zea mays subsp. mays.  
XX EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence; corn.  
OS Zea mays subsp. mays.  
XX EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140596P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-01411287P.  
PR 01-JUL-1999; 99US-01411842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 5.0%; Score 8; DB 3; Length 161;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95  
DB 97 GFDPEIAK 104  
|||||

RESULT 38  
ADX77098

ID ADX77098 standard; protein; 437 AA.

XX AC ADX77098;

XX DT 21-APR-2005 (first entry)

XX DE Plant full length insert polypeptide seqid 46464.

KW plant protectant; plant growth regulant; gene therapy; plant;  
KW recombinant DNA construct; physical array; plant breeding marker;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
KW growth rate; cell cycle pathway; disease resistance;  
KW galactomannan production; lignin production; plant growth regulator;  
KW yield; plant growth; plant development; seed oil; protein yield;

KW protein content.  
OS Unidentified.  
PN US2004034888-A1.  
XX  
PD 19-FEB-2004.  
XX  
XX 28-APR-2003; 2003US-00425114.  
XX  
XX 06-MAY-1999; 99US-00304517.  
PR 05-NOV-2001; 2001US-00985678.  
XX  
XX (LIUJ/) LIU J.  
PA (ZHOU/) ZHOU Y.  
PA (KOVA/) KOVALIC D K.  
PA (SCRE/) SCREEN S E.  
PA (TABAS/) TABASKA J E.  
PA (CAOY/) CAO Y.  
XX  
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;  
XX WPI; 2004-180133/17.  
DR  
XX  
XX New recombinant DNA construct, useful for improving plant tolerance to  
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
PT pests, for conferring increased resistance to plant disease, or for  
PT improving yield.  
XX  
XX Claim 1; SEQ ID NO 46464; 15pp; English.  
XX  
XX The invention describes a recombinant DNA construct comprising a  
CC polynucleotide consisting of a sequence encoding an amino acid sequence  
CC available in electronic form from the US patent office at  
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide  
CC of the invention are also useful in physical arrays of molecules and as  
CC plant breeding markers. The recombinant DNA construct is useful for  
CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, for conferring  
CC increased resistance to plant disease, for producing galactomannan,  
CC lignin or plant growth regulators, for increasing the rate of homologous  
CC recombination in plants, for improving yield by modification of  
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
CC or by providing improved plant growth and development under at least one  
CC stress condition or for modifying seed oil or protein yield and/or  
CC content. This is the amino acid sequence of a plant full length insert  
CC polypeptide that can be used in the recombinant DNA construct of the  
CC invention.  
XX  
SQ Sequence 437 AA;  
Query Match 5.0%; Score 8; DB 8; Length 437;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 88 GFDPEIAK 95  
DB 160 GFDPEIAK 167  
RESULT 39  
ADN19962  
ID ADN19962 standard; protein; 478 AA.  
XX  
AC ADN19962;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #2615.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW

KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
XX US2003233675-A1.  
PN  
XX  
PD 18-DEC-2003.  
XX  
XX 20-FEB-2003; 2003US-00369493.  
PF  
XX  
XX 21-FEB-2002; 2002US-0360039P.  
PR  
XX  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
PI  
XX WPI; 2004-061375/06.  
DR  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 2615; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at segdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 478 AA;  
Query Match 5.0%; Score 8; DB 8; Length 478;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 88 GFDPEIAK 95  
DB 219 GFDPEIAK 226  
RESULT 40  
ADS43079  
ID ADS43079 standard; protein; 910 AA.  
XX  
XX ADS43079;  
AC



Search completed: January 10, 2006, 13:39:47  
Job time : 78 secs

XX 02-DEC-2004 (first entry)  
XX DT  
XX DE  
XX DE Bacterial polypeptide #21509.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
XX Bacteria.  
OS  
XX US2003233675-A1.  
PN  
XX 18-DEC-2003.  
PD  
XX 20-FEB-2003; 2003US-00369493.  
PF  
XX 21-FEB-2002; 2002US-0360039P.  
PR  
XX (CAOY/) CAO Y.  
XX (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
DR  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 21509; 122pp; English.  
PS  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
XX Sequence 910 AA;  
SQ  
Query Match 5.0%; Score 8; DB 8; Length 910;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 147 AEAELKAM 154  
Db 424 AEAELKAM 431

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:36:51 ; Search time 16 Seconds

(without alignments)

956.154 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHQNLSQFFIEHILQI.....QVDGKVAEAEELKAMIAERD 159

Scoring table: OLIGO 1  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size: 8

Total number of hits satisfying chosen parameters: 18

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database: PIR 80.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	159	100.0		159	2	H64691	(3R)-hydroxymyrist
2	118	74.2		159	2	H71826	(3R)-hydroxymyrist
3	12	7.5		140	2	B95049	hypothetical prote
4	12	7.5		140	2	H97919	(3R)-hydroxymyrist
5	11	6.9		141	2	C97338	hydroxymyristoyl-(
6	10	6.3		146	2	H81445	(3R)-hydroxymyrist
7	9	5.7		151	2	D64742	(3R)-hydroxymyrist
8	9	5.7		151	2	F85502	(3R)-hydroxymyrist
9	9	5.7		151	2	F90651	hypothetical prote
10	9	5.7		181	2	AF0129	(3R)-hydroxymyrist
11	9	5.7		201	2	S35968	(3R)-hydroxymyrist
12	8	5.0		105	2	S55691	keratin type I (cl
13	8	5.0		137	2	B72335	(3R)-hydroxymyrist
14	8	5.0		368	2	A28825	keratin, type I no
15	8	5.0		478	2	S74603	dihydrolipoamide d
16	8	5.0		613	2	G69429	Glu-tRNA amidotran
17	8	5.0		841	2	T01011	hypothetical prote
18	8	5.0		910	2	C69069	cation-transportin

#### ALIGNMENTS

RESULT 1

H64691

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pylori

C/Species: Helicobacter pylori

C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004

C/Accession: H64691

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Surton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: H64691

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-159 <TOM>

A:Cross-references: UNIPROT:O25928; UNIPARC:UPI000012A4AD; GB:AE000637; GB:AB000511; NID:

C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 100.0%; Score 159; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 1.2e-158;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEQSHQNLSQFFIEHILQIILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDVFNHGPP 60

Db 1 MEQSHQNLSQFFIEHILQIILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDVFNHGPP 60

Qy 61 NKPIPPGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTKIVYFMTIDKVKRIPVTPGDRLE 120

Db 61 NKPIPPGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTKIVYFMTIDKVKRIPVTPGDRLE 120

Qy 121 YHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAERD 159

Db 121 YHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAERD 159

RESULT 2

H71826

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pylori

C/Species: Helicobacter pylori

A:Variety: strain J99

C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C/Accession: H71826

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: H71826

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-159 <ARN>

A:Cross-references: UNIPROT:Q9ZJL6; UNIPARC:UPI000012A4AC; GB:AE001551; GB:AB001439; NID:

A:Experimental source: strain J99

C:Genetics:

A:Gene: fabZ

C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C:Keywords: carbon-oxygen lyase; hydro-lyase; hydrolase

Query Match 74.2%; Score 118; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 9.9e-116;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 KIVAYKNITFNEDVFNHGPPNKPIPPGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTKIVY 100

Db 41 KIVAYKNITFNEDVFNHGPPNKPIPPGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTKIVY 100

Qy 101 FMTIDKVKFRIPVTPGDRLEHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAER 158

Db 101 FMTIDKVKFRIPVTPGDRLEHLEVLKHKGMIVQVGGTAQVDGKVAEAEELKAMIAER 158

RESULT 3

B95049

hypothetical protein SP0424 [imported] - Streptococcus pneumoniae (strain TIGR4)

C/Species: Streptococcus pneumoniae

C/Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C;Accession: B95049  
R;Tetcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eissen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A;Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: B95049  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-140 <KUR>  
A;Cross-references: UNIPROT:P59201; UNIPARC:UPI0000051423; GB:AE005672; PIDN:AAK74587.1;  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0424  
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

Query Match 7.5%; Score 12; DB 2; Length 140;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
|||||  
Db 11 LPHRYPMLLVDR 22

RESULT 4  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) [imported] - Streptococcus pneumoniae  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: H97919  
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A;Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: H97919  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-140 <KUR>  
A;Cross-references: UNIPROT:P59202; UNIPARC:UPI000012A4BA; GB:AE007317; PIDN:AAK99188.1;  
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C;Genetics:  
A;Gene: fabZ  
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 7.5%; Score 12; DB 2; Length 140;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
|||||  
Db 11 LPHRYPMLLVDR 22

RESULT 5  
hydroxymyristoyl-[acyl carrier protein] dehydratase [imported] - Clostridium acetobutyli  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: C97338  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clo*  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: C97338  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-141 <KUR>  
A;Cross-references: UNIPROT:Q97DA9; UNIPARC:UPI0000012A4A6; GB:AE001437; PIDN:AAK81494.1;  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC3571  
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

Query Match 6.9%; Score 11; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 0.00097;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PHRYPMLLVDR 32  
|||||  
Db 13 PHRYPMLLVDR 23

RESULT 6  
H81445  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) Cj0273 [imported] -  
C;Species: Campylobacter jejuni  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C;Accession: H81445  
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A;Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp  
A;Reference number: A81250; MUID:20150912; PMID:10688204  
A;Accession: H81445  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-146 <PAR>  
A;Cross-references: UNIPROT:Q9FIM2; UNIPARC:UPI0000012A4A2; GB:AL139074; GB:AL111168; NID:  
A;Experimental source: serotype O2, strain NCTC 11168  
C;Genetics:  
A;Gene: fabZ; Cj0273  
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 6.3%; Score 10; DB 2; Length 146;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEKAKMI 155  
|||||  
Db 134 VAEAEKAKMI 143

RESULT 7  
D64742  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Escherichia coli  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C;Accession: D64742; B33171  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of *Escherichia coli* K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: D64742  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-151 <BLAT>  
A;Cross-references: UNIPROT:P21774; UNIPARC:UPI0000012A4A9; GB:AE000127; GB:U00096; NID:G  
A;Experimental source: strain K-12, substrain MG1655  
R;Coleman, J.; Raetz, C.R.H.  
J. Bacteriol. 170, 1268-1274, 1988  
A;Title: First committed step of lipid A biosynthesis in *Escherichia coli*: sequence of th  
A;Reference number: A33171; MUID:88139188; PMID:3277952  
A;Accession: B33171  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-139, 'DR', 142-149, 'RRDT', <COL>  
A;Cross-references: UNIPARC:UPI0000178D50; GB:M19334; GB:M18265; GB:M18266

C;Genetics:

A;Gene: fabZ

A;Start codon: TTG

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 5.7%; Score 9; DB 2; Length 151;

Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLI 70

Db 58 KPIPPGVLI 66

RESULT 8

F85502

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) [similarity] - Esc

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C;Accession: F85502

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: F85502

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-151 &lt;STO&gt;

A;Cross-references: UNIPROT:P21774; UNIPARC:UPI000012A4A9; GB:AE005174; NID:gl2512908; E

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: fabZ

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match

Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLI 70

Db 58 KPIPPGVLI 66

RESULT 9

F90651

hypothetical protein ECs0182 [imported] - Escherichia coli (strain O157:H7, substrain R)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004

C;Accession: F90651

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90651

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-151 &lt;HAY&gt;

A;Cross-references: UNIPROT:P21774; UNIPARC:UPI00001653A3; GB:BA000007; PIDN:BA033605.1;

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs0182

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

Query Match

Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLI 70

Db 58 KPIPPGVLI 66

Db 58 KPIPPGVLI 66

RESULT 10

AF0129

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) (EC 4.2.1.-) [import

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AF0129

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; E

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell, E

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AF0129

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-181 &lt;KUR&gt;

A;Cross-references: UNIPROT:Q8ZH57; UNIPARC:UPI000012A4C3; GB:AL590842; PIDN:CAC89897.1;

C;Genetics:

A;Gene: fabZ

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match

Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLI 70

Db 58 KPIPPGVLI 66

RESULT 11

S35968

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Yersinia enteroc

C;Species: Yersinia enterocolitica

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C;Accession: S35968

R;Vuorio, R.; Harkonen, T.; Vaara, M.

submitted to the EMBL Data Library, August 1993

A;Reference number: S35966

A;Accession: S35968

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-201 &lt;VUU&gt;

A;Cross-references: UNIPROT:P32205; UNIPARC:UPI00001705F3; EMBL:Z25463; NID:g397536; PIDN

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match

Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLI 70

Db 81 KPIPPGVLI 89

RESULT 12

S55691

keratin type I (clone Jn7-4) - African clawed frog (fragment)

C;Species: Xenopus laevis (African clawed frog)

C;Date: 28-Oct-1995 #sequence\_revision 07-Feb-1997 #text\_change 09-Jul-2004

C;Accession: S55691

R;Adati, N.; Ito, T.; Koga, C.; Kito, K.; Sakaki, Y.; Shiokawa, K.

Biochim. Biophys. Acta 1262, 43-51, 1995

A;Title: Differential display analysis of gene expression in developing embryos of Xenopi

A;Reference number: S55691; MUID:95290493; PMID:7772598

A;Accession: S55691

A;Molecule type: mRNA

A;Residues: 1-105 &lt;ADA&gt;

A; Cross-references: UNIPROT:Q91404; UNIPARC:UPI00000FC2C7; EMBL:S78089; NID:g999222; PDB:1YK6  
C; Superfamily: cytoskeletal keratin  
C; Keywords: coiled coil; intermediate filament

Query Match 5.0%; Score 8; DB 2; Length 105;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 VDGKVVAE 148  
| | | | |  
Db 89 VDGKVVAE 96

RESULT 13  
B72335  
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Thermotoga maritima  
C; Species: Thermotoga maritima  
C; Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C; Accession: B72335  
R; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D., et al.  
C; M.

Nature 399, 323-329, 1999  
A; Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing  
A; Reference number: A72200; MUID:99287316; PMID:10360571  
A; Accession: B72335  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-137 <ARN>  
A; Cross-references: UNIPROT:Q9WZQ8; UNIPARC:UPI000012M4BC; GB:AE001747; GB:AE0000512; NID:  
A; Experimental source: strain MSB8  
C; Genetics:

A; Gene: TM0801  
C; Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
C; Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 5.0%; Score 8; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 PGVLIVEG 73  
| | | | |  
Db 54 PGVLIVEG 61

RESULT 14  
A28825  
keratin, type I nonepidermal - African clawed frog  
C; Species: Xenopus laevis (African clawed frog)  
C; Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C; Accession: A28825  
R; LaFlamme, S.B.; Jamrich, M.; Richter, K.; Sargent, T.D.; Dawid, I.B.  
Genes Dev. 2, 853-862, 1988  
A; Title: Xenopus endo B is a keratin preferentially expressed in the embryonic notochord  
A; Reference number: A28825; MUID:89092007; PMID:2463213  
A; Accession: A28825  
A; Molecule type: mRNA  
A; Residues: 1-368 <LAP>  
A; Cross-references: UNIPROT:P08802; UNIPARC:UPI000012DAEC; GB:Y00230; NID:g64863; PIDN:  
C; Genetics:  
A; Start codon: GGT  
C; Superfamily: cytoskeletal keratin  
C; Keywords: coiled coil

Query Match 5.0%; Score 8; DB 2; Length 368;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 VDGKVVAE 148  
| | | | |  
Db 352 VDGKVVAE 359

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, November 1997  
A;Description: Arabidopsis thaliana chromosome II BAC T517 genomic sequence.  
A;Reference number: Z14162  
A;Accession: T01011  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-841 <ROU>  
A;Cross-references: UNIPROT:O22289; UNIPARC:UPI000000AB839; EMBL:AC003000; NID:g2642152;  
A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: G84821  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-841 <STO>  
A;Cross-references: UNIPARC:UPI000000AB839; GB:AE002093; NID:g2642163; PIDN:AAB87130.1; G  
C;Genetics:  
A;Gene: T517.11; At2g39810  
A;Map position: 2  
A;Introns: 133/2; 200/3; 238/3; 365/3; 529/3; 575/3; 788/2  
C;Superfamily: Arabidopsis thaliana hypothetical protein At2g39810

Query Match 5.0%; Score 8; DB 2; Length 841;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 LVDRIIEL 36  
Db 486 LVDRIIEL 493

RESULT 18  
C69069  
cation-transporting P-ATPase PacL - Methanobacterium thermoautotrophicum (strain Delta H  
C;Species: Methanobacterium thermoautotrophicum  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 31-Dec-2004  
C;Accession: C69069  
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.  
Xi, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A;Reference number: A69000; MUID:98037514; PMID:9371463  
A;Accession: C69069  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-910 <MTH>  
A;Cross-references: UNIPROT:O27560; UNIPARC:UPI0000066763; GB:AE000912; GB:AE000666; NID  
A;Experimental source: strain Delta H  
C;Genetics:  
A;Gene: MTH1516  
A;Start codon: GTG  
C;Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom  
F;533-704/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 5.0%; Score 8; DB 2; Length 910;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 AEAELKAM 154  
Db 424 AEAELKAM 431

Search completed: January 10, 2006, 13:41:35  
Job time : 16 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:35:56 ; Search time 71 Seconds

(without alignments)  
1579.986 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHQLQSQFFIEHLQI.....QVDGKVVAEELKAMIAERD 159

Scoring table: OLIGO 60.0  
Gapop 60.0 , Gapext 60.0

Searched: 2166443 seqs, 705528306 residues

Word size : 8

Total number of hits satisfying chosen parameters: 70

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Uniprot 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	100.0	159	1	FABZ HELPY
2	118	74.2	159	1	FABZ HELPY
3	81	50.9	159	2	Q5G940 HELPY
4	14	8.8	169	1	FABZ HELPY
5	13	8.2	133	2	Q9FAV8_9STPE
6	13	8.2	142	1	FABZ CLOTE
7	12	7.5	139	1	FABZ STRP3
8	12	7.5	139	1	FABZ STRP6
9	12	7.5	139	1	FABZ STRP8
10	12	7.5	139	1	FABZ STRPY
11	12	7.5	140	1	FABZ STRA3
12	12	7.5	140	1	FABZ STRA5
13	12	7.5	140	1	FABZ STRPN
14	12	7.5	140	1	FABZ STRR6
15	12	7.5	141	1	FABZ STRT1
16	12	7.5	141	1	FABZ STRT2
17	12	7.5	164	1	FABZ RALSO
18	11	6.9	116	2	Q7PA53_FUSNV
19	11	6.9	141	1	FABZ CLOAB
20	10	6.3	142	2	Q4HG66_CAMCO
21	10	6.3	142	2	Q4HJPS_CAMLA
22	10	6.3	146	1	FABZ CAMJE
23	10	6.3	146	1	FABZ CAMJR
24	10	6.3	146	2	Q4HQ33_CAMUP
25	9	5.7	141	1	FABZ THETN
26	9	5.7	150	1	FABZ PHOLL
27	9	5.7	151	1	FABZ BORBR
28	9	5.7	151	1	FABZ BORPA
29	9	5.7	151	1	FABZ BORPE
30	9	5.7	151	1	FABZ CANBIF
31	9	5.7	151	1	FABZ_ECO57
					P0A6G8 escherichia

32	9	5.7	151	1	FABZ_ECOL6
33	9	5.7	151	1	FABZ_ECOLI
34	9	5.7	151	1	FABZ_ERWCT
35	9	5.7	151	1	FABZ_SHIFL
36	9	5.7	151	1	FABZ_WOLSU
37	9	5.7	176	1	FABZ_YERPS
38	9	5.7	178	1	FABZ_YEREN
39	9	5.7	181	1	FABZ_YERPE
40	9	5.7	298	2	Q8Y0Y7_RALSO
41	8	5.0	51	2	Q4YGS7_PLABE
42	8	5.0	94	2	Q90X75_GILMI
43	8	5.0	105	2	Q91404_XENLA
44	8	5.0	137	1	FABZ_THENA
45	8	5.0	142	1	FABZ_SYMTH
46	8	5.0	150	2	Q5E3F0_VIBF1
47	8	5.0	151	1	FABZ_IDILO
48	8	5.0	152	1	FABZ_PASMU
49	8	5.0	154	1	FABZ_RHIME
50	8	5.0	159	2	Q9T1E3_9CAUD
51	8	5.0	229	2	Q745H9_MYCPA
52	8	5.0	326	2	Q6M145_METMP
53	8	5.0	368	1	KIC18_XENLA
54	8	5.0	394	2	Q518P3_ENTHI
55	8	5.0	428	2	Q6INH6_XENLA
56	8	5.0	432	2	Q7SY65_XENLA
57	8	5.0	436	2	Q8AV12_XENLA
58	8	5.0	473	1	DLDH_SYNY3
59	8	5.0	561	2	Q94CN9_ORYSA
60	8	5.0	590	2	Q75M05_ORYSA
61	8	5.0	595	2	Q8JTV1_LSDV
62	8	5.0	596	2	Q91MW7_LSDV
63	8	5.0	613	1	GATE_ARCFU
64	8	5.0	642	2	Q50VK5_ENTHI
65	8	5.0	684	2	Q82J56_STRAW
66	8	5.0	841	2	Q22289_ARATH
67	8	5.0	856	2	Q4HMA4_CAMCO
68	8	5.0	910	2	Q27560_METTH
69	8	5.0	927	2	Q84JU6_ARATH
70	8	5.0	954	2	Q54LX3_DICDI

#### ALIGNMENTS

#### RESULT 1

FABZ_HELPY	ID	FABZ_HELPY	STANDARD;	PRT;	159 AA.
AC	O25928;				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	10-MAY-2005 (Rel. 47, Last annotation update)				
DE	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)				
DE	((3R)-hydroxymyristoyl ACP dehydratase).				
GN	Name=fabz; OrderedLocusNames=HPI1376;				
OS	Helicobacter pylori (Campylobacter pylori).				
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;				
OC	Helicobacteraceae; Helicobacter.				
OX	NCBI_TaxID=210;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].				
RC	STRAIN=26695 / ATCC 700392;				
RX	MEDLINE=97394467; Pubmed=9252185; DOI=10.1038/41483;				
RA	Tomb J.-P., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,				
RA	Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,				
RA	Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,				
RA	Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,				
RA	Glodek A., McKenney K., FitzGerald L.M., Lee N., Adams M.D.,				
RA	Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,				
RA	Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,				
RA	Wattley L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,				
RA	Smith H.O., Fraser C.M., Venter J.C.;				
RT	"The complete genome sequence of the gastric pathogen Helicobacter				
RT	pylori."				



```
RL Nature 388:539-547(1997).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB000637; AAD08419.1; -; Genomic_DNA.
DR PIR; H64691; H64691.
DR TIGR; HP1376; -.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRPFAM8; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 58 By similarity.
SQ SEQUENCE 159 AA; 18196 MW; 715B6FBD872AE32 CRC64;

Query Match 100.0%; Score 159; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.9e-152; Indels 0; Gaps 0;
Matches 159; Conservative 0; Mismatches 0;

QY 1 MEQSHQNLSQFFIEHILQILPHRYPMLLVDRILIELQANKKIYAYKNITFNEDVFNGHFP 60
DB 1 MEQSHQNLSQFFIEHILQILPHRYPMLLVDRILIELQANKKIYAYKNITFNEDVFNGHFP 60
QY 61 NKPIPGVLIVGMAQTGGFLAFTSLWGFDEPEIAKTIVFMTIDKVKPRIPVTPGDRLE 120
DB 61 NKPIPGVLIVGMAQTGGFLAFTSLWGFDEPEIAKTIVFMTIDKVKPRIPVTPGDRLE 120
QY 121 YHLEVLKHGMIWVGTAQVDGKVAEAEKAMIAERD 159
DB 121 YHLEVLKHGMIWVGTAQVDGKVAEAEKAMIAERD 159

RESULT 2
FABZ_HELPJ STANDARD; PRT; 159 AA.
AC Q9ZJL6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))
GN Names-fabZ; OrderedLocusNames=JHP1290;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE-99120557; PubMed-9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
```

```
DR EMBL; AB001551; AAD06864.1; -; Genomic_DNA.
DR PIR; H71826; H71826.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRPFAM8; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 58 By similarity.
SQ SEQUENCE 159 AA; 18210 MW; C0F16FBD868AE39 CRC64;

Query Match 74.2%; Score 118; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.2e-110; Indels 0; Gaps 0;
Matches 118; Conservative 0; Mismatches 0;

QY 41 KIVAYKNITFNEDVFNGHFPNKPPIPGVLIVGMAQTGGFLAFTSLWGFDEPEIAKTIVY 100
DB 41 KIVAYKNITFNEDVFNGHFPNKPPIPGVLIVGMAQTGGFLAFTSLWGFDEPEIAKTIVY 100
QY 101 FMTIDKVKFRIPVTPGDRLEHVLKHGMIWVGTAQVDGKVAEAEKAMIAER 158
DB 101 FMTIDKVKFRIPVTPGDRLEHVLKHGMIWVGTAQVDGKVAEAEKAMIAER 158

RESULT 3
QSG940_HELPY PRELIMINARY; PRT; 159 AA.
ID QSG940;
AC QSG940;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE ((3R)-hydroxymyristoyl-acyl carrier protein dehydratase.
GN Names-fabZ;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15967411;
RX STRAIN=SS1;
RA Liu W., Luo C., Han C., Peng S., Yang Y., Yue J., Shen X., Jiang H.;
RT "A new beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from
RT Helicobacter pylori: Molecular cloning, enzymatic characterization,
RT and structural modeling.";
RL Biochem. Biophys. Res. Commun. 333:1078-1086(2005).
DR EMBL; AV725427; AAM22049.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016836; F:hydro-lyase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR010084; FabZ.
DR TIGRPFAM8; TIGR01750; fabZ; 1.
SQ SEQUENCE 159 AA; 18184 MW; B7F16AAA78315B74 CRC64;

Query Match 50.9%; Score 81; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 4.5e-73;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGFLAFTSLWGFDEPEIAKTIVFMTIDKVKFRIPVTPGDRLEHVLKHGMIWQVGG 137
DB 78 GGFLAFTSLWGFDEPEIAKTIVFMTIDKVKFRIPVTPGDRLEHVLKHGMIWQVGG 137
QY 138 TAQVDGKVAEAEKAMIAER 158
DB 138 TAQVDGKVAEAEKAMIAER 158

RESULT 4
FABZ_HELPJ STANDARD; PRT; 169 AA.
AC Q7U3L9;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))
```

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name-fabz; OrderedLocusNames=HH1181;

OS Helicobacter hepaticus.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteraceae; Helicobacter.

OX NCBI\_TaxID=32025;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=ATCC 51449 / 3B1;

RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;

RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,

RA Bell M., Droege M., Partmann B., Fischer H.-P., Ge Z., Hoerster A.,

RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,

RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;

RT "The complete genome sequence of the carcinogenic bacterium Helicobacter hepaticus.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).

CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AE01147; AAP77778.1; -; Genomic\_DNA.

DR HAMAP; MF\_00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabZ; 1.

DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

KW ACT SITE 66 By similarity.

FT ACT SITE 66

SQ SEQUENCE 169 AA; 19071 MW; 011CAB723E36E25C CRC64;

Query Match 8.8%; Score 14; DB 1; Length 169;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 QILPHRYPMMLVDVR 32

DB 20 QILPHRYPMMLVDVR 33

RESULT 5

Q9FAY8\_9STRE PRELIMINARY; PRT; 133 AA.

AC Q9FAY8;

DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)

DE Similar to Bacillus subtilis hydroxymyristoyl-(acyl carrier protein) dehydratase (Fragment).

DE Streptococcus equi subsp. zooepidemicus.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=40041;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=VTU211;

RA Hong K.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB030883; BAB16042.1; -; Genomic\_DNA.

DR HSSP; P18391; IMKA.

DR GO; GO:0005737; C:cytoplasm; IEA.

DR GO; GO:0016836; F:hydro-lyase activity; IEA.

DR GO; GO:0016829; F:lyase activity; IEA.

DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.

DR GO; GO:0009245; P:lipid A biosynthesis; IEA.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabZ; 1.

FT NON TER 1

SQ SEQUENCE 133 AA; 14665 MW; 70PBF56236598A07 CRC64;

Query Match 8.2%; Score 13; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMMLVDRI 33

DB 4 LPHRYPMMLVDRI 16

RESULT 6

FABZ\_CLOTE

ID -FABZ\_CLOTE STANDARD; PRT; 142 AA.

AC Q899N7;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.-)

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabz; OrderedLocusNames=CTC00133;

OS Clostridium tetani.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1513;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Massachusetts / E88;

RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;

RA Brueggemann H., Baeumer S., Fricke W.F., Wierzer A., Liesegang H.,

RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,

RA Gottschalk G.;

RT "The genome sequence of Clostridium tetani, the causative agent of tetanus disease.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).

CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AE015936; AAC34785.1; -; Genomic\_DNA.

DR HAMAP; MF\_00406; -; 1

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabZ; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 50 By similarity.

SQ SEQUENCE 142 AA; 15706 MW; B8D7A9F53CF336 CRC64;

Query Match 8.2%; Score 13; DB 1; Length 142;

Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ILQILPHRYPMML 29

DB 9 ILQILPHRYPMML 21

RESULT 7

FABZ\_STRP3

ID -FABZ\_STRP3 STANDARD; PRT; 139 AA.

AC Q8K631;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.-)

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabz; OrderedLocusNames=SPYM3\_1520, SPs0346;

OS Streptococcus pyogenes (serotype M3).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;



KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT ACT\_SITE 46 By similarity.  
SQ SEQUENCE 139 AA; 15324 MW; 06090333FC20B479 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 139;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
| | | | | | | | | |  
DB. 10 LPHRYPMLLVDR 21

RESULT 10  
FABZ\_STRPY STANDARD; PRT; 139 AA.  
ID FABZ\_STRPY  
AC P64110; P58175;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1-)  
DE (3R)-hydroxymyristoyl ACP dehydratase).  
GN Name-fabZ; OrderedLocusNames=SPY1746;  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;  
RX Ferrerini J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;  
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

EMBL; AE006603; AAK34490.1; -; Genomic\_DNA.  
DR HSSP; P18391; IMKA.  
DR HAMAP; MF\_00406; -; 1.  
DR InterPro; IPR010084; FabZ.  
DR TIGRFAMs; TIGR01750; fabZ; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT ACT\_SITE 46 By similarity.  
SQ SEQUENCE 139 AA; 15324 MW; 06090333FC20B479 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 139;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
| | | | | | | | | |  
DB. 10 LPHRYPMLLVDR 21

RESULT 11  
FABZ\_STR33 STANDARD; PRT; 140 AA.  
ID FABZ\_STR33  
AC Q8E720;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1-)

DE (3R)-hydroxymyristoyl ACP dehydratase).  
GN Name-fabZ; OrderedLocusNames=gbs0338;  
OS Streptococcus agalactiae (serotype III).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=216495;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=22242508 / Serotype III;  
RX MEDLINE=22242508; PubMed=12354221;  
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L.,  
Maadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,  
Kunst F.;  
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing  
RT invasive neonatal disease.";  
RL Mol. Microbiol. 45:1499-1513(2002).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

EMBL; AL766844; CAD45983.1; -; Genomic\_DNA.  
DR HSSP; P18391; IMKA.  
DR Sagaliet; gbs0338; -;  
DR HAMAP; MF\_00406; -; 1.  
DR InterPro; IPR010084; FabZ.  
DR TIGRFAMs; TIGR01750; fabZ; 1.  
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT ACT\_SITE 47 By similarity.  
SQ SEQUENCE 140 AA; 15445 MW; PF12DAED4015226B CRC64;

QY 21 LPHRYPMLLVDR 32  
| | | | | | | | | |  
DB. 11 LPHRYPMLLVDR 22

RESULT 12  
FABZ\_STRAS STANDARD; PRT; 140 AA.  
ID FABZ\_STRAS  
AC Q8E1K3;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1-)  
DE (3R)-hydroxymyristoyl ACP dehydratase).  
GN Name-fabZ; OrderedLocusNames=SAG0351;  
OS Streptococcus agalactiae (serotype V).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=216466;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=2603 V/R / ATCC BAA-611 / Serotype V;  
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;  
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,  
Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,  
Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,  
DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,  
Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,  
Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarcelli M., Mora M.,  
Rinaldo E.F., Brettoni R., Galli G., Mariani M., Vegni F., Malone D.,  
Fraser C.M.;  
RT "Complete genome sequence and comparative genomic analysis of an

RT emerging human pathogen, serotype V *Streptococcus agalactiae*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).  
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; AE014207; AAM99257.1; -; Genomic\_DNA.  
 DR HSSP; P18391; IMKA.  
 DR TIGR; SAG0351; -.  
 DR HAMAP; MF 00406; -; 1.  
 DR InterPro; IPRO10084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 47 By similarity.  
 SQ SEQUENCE 140 AA; 15445 MW; FF12DAED4015226B CRC64;  
 -----  
 Query Match 7.5%; Score 12; DB 1; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 0.0016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 QY 21 LPHRYPMLLVDR 32  
 DB 11 LPHRYPMLLVDR 22  
 |||||  
 -----  
 RESULT 13  
 FABZ\_STRPN STANDARD; PRT; 140 AA.  
 AC P59201; Q9FBC0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))  
 DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 GN Names=fabZ; OrderedLocNames=SP0424;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=ATCC BAA-334 / TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;  
 RA Tettelin H., Nelson K.E., Paulsen I.T., Haft D.H., Dodson R.J.,  
 RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,  
 RA Durkin A.S., Gwinn M.B., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,  
 RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 RT "Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*."  
 RL Science 293:498-506(2001).  
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; AE007354; AAK74587.1; -; Genomic\_DNA.  
 DR PIR; B95049; B95049.  
 DR HSSP; P18391; IMKA.

DR TIGR; SP0424; -.  
 DR HAMAP; MF 00406; -; 1.  
 DR InterPro; IPRO10084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 47 By similarity.  
 SQ SEQUENCE 140 AA; 15272 MW; A54650DDA9834813 CRC64;  
 -----  
 Query Match 7.5%; Score 12; DB 1; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 0.0016;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 QY 21 LPHRYPMLLVDR 32  
 DB 11 LPHRYPMLLVDR 22  
 |||||  
 -----  
 RESULT 14  
 FABZ\_STRR6 STANDARD; PRT; 140 AA.  
 ID P59202; Q9FBC0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))  
 DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 GN Names=fabZ; OrderedLocNames=spr0384;  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=171101;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20365714; PubMed=10910344; DOI=10.1038/35018162;  
 RA Heath R.J., Rock C.O.;  
 RT "A triclofen-resistant bacterial enzyme."  
 RL Nature 406:145-146(2000).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX MEDLINE=21429245; PubMed=11544234;  
 RX DOI=10.1128/JB.183.19.5709-5717.2001;  
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,  
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,  
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,  
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,  
 RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,  
 RA Norris P.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,  
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,  
 RA Glass J.I.;  
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6."  
 RL J. Bacteriol. 183:5709-5717(2001).  
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; AF197933; AAF98278.1; -; Genomic\_DNA.  
 DR EMBL; AE008418; AAK99188.1; -; Genomic\_DNA.  
 DR PIR; H97919; H97919.  
 DR HSSP; P18391; IMKA.  
 DR HAMAP; MF 00406; -; 1.  
 DR InterPro; IPRO10084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 47 By similarity.  
 FT CONFLICT 58 L -> V (in Ref. 1).  
 FT CONFLICT 133 I -> T (in Ref. 1).

```

SQ SEQUENCE 140 AA; 15298 MW; 6646401713F382B8 CRC64;
Query Match 7.5%; Score 12; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 22

RESULT 15
FABZ_STRT1
ID FABZ_STRT1 STANDARD; PRT; 141 AA.
AC QSM177;
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=stu0390;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fongstein M., Overbeek R., Kyprides N., Fumelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtan S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CP000024; AAV61992.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 141 AA; 15445 MW; 66E18ED333642A67 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 23

RESULT 16
FABZ_STRT2
ID FABZ_STRT2 STANDARD; PRT; 141 AA.
AC QSM5R4;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=stu0390;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fongstein M., Overbeek R., Kyprides N., Fumelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtan S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CP000024; AAV61992.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 141 AA; 15445 MW; 66E18ED333642A67 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 23

RESULT 17
FABZ_STRT3
ID FABZ_STRT3 STANDARD; PRT; 164 AA.
AC Q8XZ10;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=RS05277;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=GM1000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brothier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Javie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

```

OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fongstein M., Overbeek R., Kyprides N., Fumelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtan S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CP000023; AAV60108.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 141 AA; 15445 MW; 66E18ED333642A67 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 12 LPHRYPMLLVDR 23

RESULT 17
FABZ_STRT4
ID FABZ_STRT4 STANDARD; PRT; 164 AA.
AC Q8XZ10;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=RS05277;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=GM1000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brothier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Javie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```



CC between the Swiss Institute of Bioinformatics and the ENBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC -----  
ENBL; AL646064; CAD15117.1; -; Genomic\_DNA.  
DR HAMAP; MF 00406; -; 1.  
DR InterPro; IPR010084; FabZ.  
DR TIGRFAMs; TIGR01750; fabZ; 1.  
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT SITE 61 By similarity.  
ST SEQUENCE 164 AA; 19461 MW; 69FAF8137B2A9A29 CRC64;  
FO

Query Match	7.5%	Score 12;	DB 1;	Length 164;
Best Local Similarity	100.0%;	Pred. No. 0.0018;		
Matches 12;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps 0;				

**Qy** 21 LPHRYPMLLVDR 32  
||| |||  
**Db** 24 LPHRYPMLLVDR 35

RESULT 18  
Q7P453 FUSNV  
ID Q7P453 FUSNV PRELIMINARY;  
PRT: 116 AA.

AC Q7P453;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase  
 DE (EC 4.2.1.-).

OS Fusobacterium nucleatum subsp. vincentii ATCC 49356.  
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
OC Fusobacterium.  
OX NCHI TaxID=209882;

FN	[1]	RP	RC
	NUCLEOTIDE SEQUENCE.		
	STRAIN=ATCC 49256;		

RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,  
RA Haseelkorn R., Overbeek R., Kyripides N.,  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC

CC preliminary data.  
 DR EMBL; AABP01000137; EAA23396.1; -; Genomic\_DNA.  
 DR GO: 0016829; F:lyase activity; IEA.  
 DR GO: 0008610; P:lipid biosynthesis; IEA.  
 KW Lyase.

Query Match	6.9%	Score 11	DB 2	Length 116
Best Local Similarity	100.0%	Pred. No. 0.014		
Matches 11	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

Qy	PGVLIVEGMAQ	76
Dh	PGVLIVEGMAQ	42

RESULT 19	PET:	141 AA-
FABZ_CLOAB		
ID_FABZ_CLOAB	STANDARD:	

AC Q97DA9; Created  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

GN  
OS  
OC  
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
Clostridium acetobutylicum.  
Name=fab7; OrderedLocusNames=CAC3571;  
taxonomicRank=order;

OC Clostridium.  
OX NCBI\_TaxID=1488;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=ATCC 824 / DSM 792 / VKM B-178;  
RX MEDLINE=21359325; PubMed=11466286;  
RX DOI=10.1128/JB.183.16.4823-4838.2001;  
RX Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,  
RX Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,  
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;  
RT 'Genome sequence and comparative analysis of the solvent-producing  
RT

J. Bacteriol. 183:4823-4838 (2001).  
 RL  
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC -----  
CC

CC	removed.
CC	-----
DR	EMBL: AE007854; AAK81494.1; -; Genomic DNA.

DR DR HAMAP; MF 00406; -; 1.  
DR PIR; C97338; C97338.  
DR InterPro; IPR010084; FabZ.  
DR TIGRFAMS; TIGR01750; FabZ; 1.  
DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
PT ACT\_SITE 49 49 By similarity.

Query Match	6.9%;	Score 11;	DB 1;	Length 141;
Best Local Similarity	100.0%;	Pred. No. 0.016;		
Matches 11;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	

Qy	22	PHRYPMLLVDR	32
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			

RESULT 20  
Q4HGA6 CAMCO Q4HGA6 CAMCO PRELIMINARY;  
ID Q4HGA6 CAMCO PRELIMINARY; PRT; 142 AA.

Q4HG66; AC 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ

OC Campylobacter; Enterobacteriaceae; Enterobacteriales;  
OS Campylobacter coli RM2228;  
OC Bacteria; proteobacteria; Epsilonproteobacteria; Campylobacterales;  
OC Bacteria; proteobacteria; Epsilonproteobacteria; Campylobacter.  
OC Campylobacteraceae; Campylobacter.

OR	NCBI accession number
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=RM2228;

RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,  
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,  
RA Nelson K.E.:

RT "Major structural and novel potential virulence mechanisms from the  
RT genomes of multiple Campylobacter species."; [EMBL/GenBank/DBJ](#)  
RL Submitted (DEC-2004) to the [EMBL/GenBank/DBJ](#) databases.  
CC -i- CAUTION: The sequence shown here is derived from an  
CC [EMBL/GenBank/DBJ](#) whole genome shotgun (WGS) entry which is

DR EMBL; AAFL01000003; EAL56875.1; -; Genomic\_DNA.  
KW Lyase.  
SQ SEQUENCE 142 AA; 15937 MW; 2E000F62AC59E479 CRC64;



```

Query Match          6.3%; Score 10; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEIKAMI 155
DB 130 VAEAEIKAMI 139

RESULT 21
Q4HJP5 CAMLA
ID Q4HJP5_CAMLA PRELIMINARY; PRT; 142 AA.
AC Q4HJP5;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
  (EC 4.2.1.-)
GN Name=fabZ; ORFNames=CLA0287;
OS Campylobacter lari RM2100.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306263;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2100;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
  Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
  Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
  Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
  Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
  genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
  preliminary data.
CC EMBL; AAFK01000004; EAL54516.1; -; Genomic_DNA.
DR EMBL; AAFK01000004; EAL54516.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 142 AA; 16066 MW; D4DIDBA5660BC286 CRC64;

Query Match          6.3%; Score 10; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEIKAMI 155
DB 130 VAEAEIKAMI 139

RESULT 22
FABZ CAMJE
ID FABZ_CAMJE STANDARD; PRT; 146 AA.
AC Q9PIM2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
  ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=Cj0273;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NTC 11168;
EX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RA Parkhill J., Wren B.W., Mungall K.L., Kettle J.M., Churcher C.M.,
  Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
  Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
  Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
  Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
  reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use as long as its content is in no way modified and this statement is not
  removed.
DR EMBL; AL139074; CAB72741.1; -; Genomic_DNA.
DR PIR; H81445; H81445.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT SITE 48
  By similarity.
SQ SEQUENCE 146 AA; 16440 MW; EPFA2656B12563A5 CRC64;

Query Match          6.3%; Score 10; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEIKAMI 155
DB 134 VAEAEIKAMI 143

RESULT 23
FABZ CAMJR
ID FABZ_CAMJR STANDARD; PRT; 146 AA.
AC QSHWJ3;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
  ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=CJ50322;
OS Campylobacter jejuni (strain RM1221).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=195099;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15660156; DOI=10.1371/journal.pbio.0030015;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
  Ravel J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
  Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
  Ayodeji M.A., Shvartsbeyn A., Schatz M.C., Badger J.H., Fraser C.M.,
  Nelson K.E.;
RT "Major structural differences and novel potential virulence mechanisms
  from the genomes of multiple Campylobacter species.";
RL PLOS Biol. 3:72-85(2005).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use as long as its content is in no way modified and this statement is not
  removed.
DR EMBL; CP000025; AAW34912.1; -; Genomic_DNA.
DR TIGR; CJE0322; -; 1.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

```

```

FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 146 AA; 16440 MW; EFA2656B12563A5 CRC64;

Query Match
Best Local Similarity 6.3%; Score 10; DB 1; Length 146;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEKAMI 155
Db 134 VAEAEKAMI 143

RESULT 24
ID Q4HOA3 CAMUP PRELIMINARY; PRT; 146 AA.
AC Q4HOA3;
DT 13-SEP-2005 (TREMELrel. 31, Created)
DT 13-SEP-2005 (TREMELrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMELrel. 31, Last annotation update)
DE Beta-hydroxyacyl- (Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-).
GN Name=fabZ; ORFNames=CUP1147;
OS Campylobacter upsaliensis RM3195.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306264;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM3195;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data
CC EMBL; AAFJ01000007; EAL52872.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 146 AA; 16591 MW; 8F1CD2BA02891959 CRC64;

Query Match
Best Local Similarity 6.3%; Score 10; DB 2; Length 146;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEKAMI 155
Db 134 VAEAEKAMI 143

RESULT 25
FABZ_THETN
ID FABZ_THETN STANDARD; PRT; 141 AA.
AC Q8RD71;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; Synonyms=fabA; OrderedLocusNames=TTE0175;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,

```

```

RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE012991; AAM23476.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 47 47 By similarity.
SQ SEQUENCE 141 AA; 15604 MW; 42ASA71CFDF0DBAA CRC64;

Query Match
Best Local Similarity 5.7%; Score 9; DB 1; Length 141;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LLVDRIIEL 36
Db 17 LLVDRIIEL 25

RESULT 26
FABZ_PHOLL
ID FABZ_PHOLL STANDARD; PRT; 150 AA.
AC Q7N8N6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=plu0683;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taouit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Prevassinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Sigulier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX571861; CAE12978.1; -; Genomic_DNA.
DR PhotocList; plu0683; -; 1.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 53 53 By similarity.

```

SQ SEQUENCE 150 AA; 17079 MW; 6BAF16D204B1F708 CRC64;  
 Query Match 5.7%; Score 9; DB 1; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 62 KP1PPGVL 70  
 |||||  
 DB 57 KP1PPGVL 65  
 |||||  
 RESULT 27  
 FABZ BORBR STANDARD; PRT; 151 AA.  
 AC Q7WJ83;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.1-)  
 DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 DE Name=fabZ; OrderedLocusNames=BB2616;  
 DE Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
 OS Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica."  
 RL Nat. Genet. 35:32-40(2003).  
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC EMBL; BX640427; CAE3109.1; -; Genomic\_DNA.  
 DR HAMAP; MF\_00406; -; 1.  
 DR InterPro; IPR010084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 49 By similarity.  
 SQ SEQUENCE 151 AA; 16754 MW; D992AE319C86399E CRC64;  
 Query Match 5.7%; Score 9; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 21 LPHRYPM 29  
 |||||  
 DB 12 LPHRYPM 20  
 |||||  
 RESULT 28  
 FABZ BORBR STANDARD; PRT; 151 AA.  
 AC Q7WJ83;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.1-)  
 DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 DE Name=fabZ; OrderedLocusNames=BB2616;  
 DE Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
 OS Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica."  
 RL Nat. Genet. 35:32-40(2003).  
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC EMBL; BX640444; CAE3109.1; -; Genomic\_DNA.  
 DR HAMAP; MF\_00406; -; 1.  
 DR InterPro; IPR010084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 49 By similarity.  
 SQ SEQUENCE 151 AA; 16754 MW; D992AE319C86399E CRC64;  
 Query Match 5.7%; Score 9; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 21 LPHRYPM 29  
 |||||  
 DB 12 LPHRYPM 20  
 |||||  
 RESULT 29  
 FABZ BORBR STANDARD; PRT; 151 AA.  
 AC Q7WJ83;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.1-)  
 DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 DE Name=fabZ; OrderedLocusNames=BB2616;  
 DE Bordetella pertussis.  
 OS Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=520;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA

SQ SEQUENCE 150 AA; 17079 MW; 6BAF16D204B1F708 CRC64;  
 Query Match 5.7%; Score 9; DB 1; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 62 KPIPPGVLI 70  
 |||||  
 DB 57 KPIPPGVLI 65  
 |||||  
 RESULT 27  
 FABZ BORBR STANDARD; PRT; 151 AA.  
 AC Q7WJ83;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.1.-)  
 DE ((3R)-hydroxymyristoyl-ACP dehydratase).  
 DE Name=fabZ; OrderedLocusNames=BB2616;  
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
 GN Bordetella bronchiseptica  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica."  
 RL Nat. Genet. 35:32-40(2003).  
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC EMBL; BX640427; CAE3109.1; -; Genomic\_DNA.  
 DR HAMAP; MF\_00406; -; 1.  
 DR InterPro; IPR010084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 49 By similarity.  
 SQ SEQUENCE 151 AA; 16754 MW; D992AE319C86399E CRC64;  
 Query Match 5.7%; Score 9; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 21 LPHRYPMML 29  
 |||||  
 DB 12 LPHRYPMML 20  
 |||||  
 RESULT 28  
 FABZ BORBR STANDARD; PRT; 151 AA.  
 AC Q7WJ83;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.1.-)  
 DE ((3R)-hydroxymyristoyl-ACP dehydratase).  
 DE Name=fabZ; OrderedLocusNames=BB2616;  
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
 GN Bordetella bronchiseptica  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica."  
 RL Nat. Genet. 35:32-40(2003).  
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC EMBL; BX640444; CAE3109.1; -; Genomic\_DNA.  
 DR HAMAP; MF\_00406; -; 1.  
 DR InterPro; IPR010084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 49 By similarity.  
 SQ SEQUENCE 151 AA; 16754 MW; D992AE319C86399E CRC64;  
 Query Match 5.7%; Score 9; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 21 LPHRYPMML 29  
 |||||  
 DB 12 LPHRYPMML 20  
 |||||  
 RESULT 28  
 FABZ BORBR STANDARD; PRT; 151 AA.  
 AC Q7WJ83;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.1.-)  
 DE ((3R)-hydroxymyristoyl-ACP dehydratase).  
 DE Name=fabZ; OrderedLocusNames=BB2616;  
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
 GN Bordetella bronchiseptica  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica."  
 RL Nat. Genet. 35:32-40(2003).  
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC EMBL; BX640444; CAE3109.1; -; Genomic\_DNA.  
 DR HAMAP; MF\_00406; -; 1.  
 DR InterPro; IPR010084; FabZ.  
 DR TIGRFAMs; TIGR01750; fabZ; 1.  
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 FT ACT SITE 49 By similarity.  
 SQ SEQUENCE 151 AA; 16754 MW; D992AE319C86399E CRC64;  
 Query Match 5.7%; Score 9; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 21 LPHRYPMML 29  
 |||||  
 DB 12 LPHRYPMML 20  
 |||||

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
RA Achman M., Atkin R., Baker S., Basham K., Bason N., Cherevach I.,  
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,  
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
RT Bordetella parapertussis and Bordetella bronchiseptica.";  
RL Nat. Genet. 35:32-40(2003).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; BX640415; CAB41720.1; -; Genomic\_DNA.  
DR HAMAP; MF\_00406; -; 1.  
DR InterPro; TIGR010084; FabZ.  
DR TIGRFAMS; TIGR01750; fabZ, 1.  
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT ACT\_SITE 49 49 By similarity.  
SQ SEQUENCE 151 AA; 16754 MW; D992AE319C86399E CRC64;  
  
Query Match 5.7%; Score 9; DB 1; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 21 LPHRYPMILL 29  
DB 12 LPHRYPMILL 20  
|||||  
|  
  
RESULT 30  
FABZ\_CANBF STANDARD; PRT; 151 AA.  
ID FABZ\_CANBF  
AC Q7VRD5;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE ((3R)-hydroxymyristoyl-lacyl carrier protein) dehydratase (EC 4.2.1.-)  
DE ((3R)-hydroxymyristoyl ACP dehydratase).  
GN Names=fabZ; OrderedLocNames=Bfl1282;  
OS Candidatus Blochmannia floridanus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.  
OX NCBI\_TaxID=203907;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=22784745; PubMed=12886019; DOI=10.1073/pnas.1533499100;  
RA Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,  
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,  
RA van Ham R.C.H.J., Gross R., Moya A.;  
RT "The genome sequence of Blochmannia floridanus: comparative analysis  
RT of reduced genomes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393(2003).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; BX248585; CAD83353.1; -; Genomic\_DNA.

DR HAMAP; MF\_00406; -; 1.  
DR InterPro; TIGR010084; FabZ.  
DR TIGRFAMS; TIGR01750; fabZ, 1.  
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT ACT\_SITE 54 54 By similarity.  
SQ SEQUENCE 151 AA; 17145 MW; 6C94FF8F89ABED3 CRC64;  
  
Query Match 5.7%; Score 9; DB 1; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 62 KPIPPGVLI 70  
DB 58 KPIPPGVLI 66  
|||||  
|  
  
RESULT 31  
FABZ\_ECO57 STANDARD; PRT; 151 AA.  
ID FABZ\_ECO57  
AC POA608; P21774;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE ((3R)-hydroxymyristoyl-lacyl carrier protein) dehydratase (EC 4.2.1.-)  
DE ((3R)-hydroxymyristoyl ACP dehydratase).  
GN Names=fabZ; OrderedLocNames=z0192, ECs0182;  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;  
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
RA Kuwara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBUNIT: Oligomer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- PTM: The N-terminus is blocked (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; AB005174; AAGS4482.1; -; Genomic\_DNA.  
DR EMBL; BA000007; BAB33605.1; -; Genomic\_DNA.  
DR PIR; F85502; F85502.  
DR PIR; F90651; F90651.  
DR HAMAP; MF\_00406; -; 1.  
DR InterPro; TIGR010084; FabZ.  
DR TIGRFAMS; TIGR01750; fabZ, 1.  
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

```

FT ACT SITE 54 54 BY similarity.
FT CONFLICT 37 37 F -> L (in Ref. 2).
SQ SEQUENCE 151 AA; 17033 MW; 91F514A7319C2FB1 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 32
FABZ_ECOLI6
ID FABZ_ECOLI6 STANDARD; PRT; 151 AA.
AC POA6Q7; P21774;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE (3R)-hydroxymyristoyl ACP dehydratase.
GN Names=fabZ; OrderedLocusNames=c0217;
OS Escherichia coli O6
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=O6:HI / CFT073 / ATCC 700928 / UPEC;
RC MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBUNIT: Oligomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- PTM: The N-terminus is blocked (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
DR EMBL; AR016755; AAN78709.1; ALT_INIT; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRfam; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT SITE 54 54 By similarity.
SQ SEQUENCE 151 AA; 17033 MW; 91F514A7319C2FB1 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 33
FABZ_ECOLI
ID FABZ_ECOLI STANDARD; PRT; 151 AA.
AC POA6Q6; P21774;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE (3R)-hydroxymyristoyl ACP dehydratase.
GN Names=fabZ; OrderedLocusNames=b0180;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE.
RP Coleman J.; Raetz C.R.H.;
RC MEDLINE=88139188; PubMed=3277952;
RA Coleman J., Raetz C.R.H.;
RT "First committed step of lipid A biosynthesis in Escherichia coli:
RT sequence of the lpxA gene.";
RL J. Bacteriol. 170:1268-1274(1988).
RN NUCLEOTIDE SEQUENCE.
RP Coleman J.;
RT Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=K12 / W3110;
RC Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=K12 / MG1655;
RC Chung E., Allen E., Araujo R., Aparicio A.M., Davis K., Duncan M.,
RA Federpriel N., Hyman R., Kallman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RT "Sequence of minutes 4-25 of Escherichia coli.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=K12 / MG1655;
RC Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN PROTEIN SEQUENCE OF 2-18, SIMILARITY, AND POSSIBLE FUNCTION.
RP STRAIN=K12 / W3110;
RC MEDLINE=93359166; PubMed=8354462; DOI=10.1016/0378-1097(93)90118-L;
RA Foster S.J.;
RT "Purification and characterization of an 'actomyosin' complex from
RT Escherichia coli W3110.";
RL FEMS Microbiol. Lett. 110:295-298(1993).
RN FUNCTION.
RP MEDLINE=95105173; PubMed=7806516;
RA Mohan S., Kelly T.M., Eveland S.S., Raetz C.R.H., Anderson M.S.;
RT "An Escherichia coli gene (fabZ) encoding (3R)-hydroxymyristoyl acyl
RT carrier protein dehydratase. Relation to fabA and suppression of
RT mutations in lipid A biosynthesis.";
RL J. Biol. Chem. 269:32896-32903(1994).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBUNIT: Oligomer.
CC -!- INTERACTION:
CC POA6A8:acpP; NbExp=1; IntAct=EBI-370073; EBI-542566;
CC P3659:cbpA; NbExp=1; IntAct=EBI-370073; EBI-546131;
CC P15716:clpA; NbExp=1; IntAct=EBI-370073; EBI-546140;
CC POA6P9:eno; NbExp=1; IntAct=EBI-370073; EBI-368855;
CC P07019:iffh; NbExp=1; IntAct=EBI-370073; EBI-369938;
CC P00962:glns; NbExp=1; IntAct=EBI-370073; EBI-550012;
CC P08177:lon; NbExp=1; IntAct=EBI-370073; EBI-547203;
CC P21513:rne; NbExp=1; IntAct=EBI-370073; EBI-549958;

```

```

CC POA7L0:rp1A; NbExp=1; IntAct=EBI-370073, EBI-543771;
CC P60422:rp1B; NbExp=1; IntAct=EBI-370073, EBI-543515;
CC P60438:rp1C; NbExp=1; IntAct=EBI-370073, EBI-542200;
CC P60723:rp1D; NbExp=1; IntAct=EBI-370073, EBI-545597;
CC P63299:rp1E; NbExp=1; IntAct=EBI-370073, EBI-545956;
CC P63390:rp1F; NbExp=1; IntAct=EBI-370073, EBI-546389;
CC POA7R1:rp1I; NbExp=1; IntAct=EBI-370073, EBI-546437;
CC POA7J3:rp1J; NbExp=1; IntAct=EBI-370073, EBI-546827;
CC P02410:rp1M; NbExp=1; IntAct=EBI-370073, EBI-543801;
CC P0A7K6:rp1S; NbExp=1; IntAct=EBI-370073, EBI-543891;
CC P61175:rp1V; NbExp=1; IntAct=EBI-370073, EBI-542255;
CC P02349:rp8A; NbExp=1; IntAct=EBI-370073, EBI-546520;
CC P0A7V0:rp8B; NbExp=1; IntAct=EBI-370073, EBI-543439;
CC P0A7V3:rp8C; NbExp=1; IntAct=EBI-370073, EBI-544851;
CC P0A7V8:rp8D; NbExp=1; IntAct=EBI-370073, EBI-543939;
CC P0A7M1:rp8E; NbExp=1; IntAct=EBI-370073, EBI-543949;
CC P02359:rp8G; NbExp=1; IntAct=EBI-370073, EBI-543074;
CC P02990:tu1A; NbExp=1; IntAct=EBI-370073, EBI-543077;
CC P77349:ybeZ; NbExp=1; IntAct=EBI-370073, EBI-560255;
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M19334; AAC36917.1; -; Genomic DNA.
CC EMBL; U00096; AAC73291.1; -; Genomic DNA.
CC EMBL; D83536; BAA7855.1; -; Genomic DNA.
CC EMBL; U70214; AAB08609.1; -; Genomic DNA.
CC PIR; D64742; D64742.
CC IntAct; P0A6Q6; -.
CC EcoBASE; EBI1261; -.
CC EcoGene; EGI1284; fabZ.
CC HAMAP; MF 00406; -; 1.
CC InterPro; IPR010084; FabZ.
CC TIGRPFAMs; TIGR01750; fabZ; 1.
CC Complete proteome; Direct protein sequencing; Lipid A biosynthesis;
CC Lipid synthesis; Lyase.
CC ACT SITE 54 54 By similarity.
CC CONFLICT 140 141 EA -> DR (in Ref. 1).
CC CONFLICT 150 151 EA -> RRD (in Ref. 1).
CC SEQUENCE 151 AA; 17033 MW; 91F514A7319C2FB1 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFFPGVLI 70
DB 58 KPIFFPGVLI 66

RESULT 34
FABZ_ERWCT
ID FABZ_ERWCT STANDARD; PRT; 151 AA.
AC Q6D8D2;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Names=fabZ; OrderedLocNames=ECA1042;
OS Ewriaia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OC NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;


```

```

RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebaihia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagers K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX950851; CAG73953.1; -; Genomic DNA.
CC HAMAP; MF 00406; -; 1.
CC InterPro; IPR010084; FabZ.
CC TIGRPFAMs; TIGR01750; fabZ; 1.
CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
CC ACT SITE 54 54 By similarity.
CC SEQUENCE 151 AA; 17105 MW; 44A0F1B878069273 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFFPGVLI 70
DB 58 KPIFFPGVLI 66

RESULT 35
FABZ_SHIFL
ID FABZ_SHIFL STANDARD; PRT; 151 AA.
AC P0A6Q9; P21774;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocNames=SF0170, S0173;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OC NCBI_TaxID=623;
OX [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RC MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew M.B., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;


```



RT "Complete genome sequence and comparative genomics of Shigella  
 RL flexneri serotype 2a strain 2457T";  
 CC Infect. Immun. 71:2775-2786(2003).  
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -|- SUBUNIT: Oligomer (By similarity).  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -|- PTM: The N-terminus is blocked (By similarity).  
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL; AB005674; AAN41832.1; ALT INIT; Genomic\_DNA.  
 CC EMBL; AB016978; AAP15713.1; -; Genomic\_DNA.  
 CC HAMAP; MF 00406; -; 1.  
 CC InterPro; IPR010084; FabZ.  
 CC TIGRFAMs; TIGR01750; fabz; 1.  
 CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 KW ACT SITE 54 By similarity.  
 FT ACT SITE 54  
 SQ SEQUENCE 151 AA; 17033 MW; 91F514A7319C2PB1 CRC64;  
 CC  
 CC Query Match 5.7%; Score 9; DB 1; Length 151;  
 CC Best Local Similarity 100.0%; Pred. No. 1.8;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 62 KPFPFGVLI 70  
 CC DB 58 KPFPFGVLI 66  
 CC  
 CC RESULT 36  
 CC FABZ WOLSU  
 CC ID FABZ WOLSU STANDARD; PRT; 151 AA.  
 CC AC Q7MA52;  
 CC DT 05-JUL-2004 (Rel. 44, Created)  
 CC DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 CC DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 CC DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)  
 CC DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 CC GN Name=fabZ; OrderedLocusNames=WS0060;  
 CC OS Wolinella succinogenes.  
 CC OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 CC OC Helicobacteraceae; Wolinella.  
 CC OX NCBI\_TaxID=844;  
 CC RN [1]  
 CC RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 CC RC STRAIN=DSMZ 1740;  
 CC RX MEDLINE=22882897; PubMed=14500908; DOI=10.1073/pnas.1932838100;  
 CC RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,  
 CC RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,  
 CC RA Meyer F., Lederer H., Schuster S.C.;  
 CC RT "Complete genome sequence and analysis of Wolinella succinogenes";  
 CC Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).  
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL; BX571657; CAE09229.1; -; Genomic\_DNA.  
 CC HAMAP; MF 00406; -; 1.  
 CC InterPro; IPR010084; FabZ.  
 CC TIGRFAMs; TIGR01750; fabz; 1.  
 CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 KW ACT SITE 49 By similarity.  
 FT ACT SITE 49  
 SQ SEQUENCE 151 AA; 17084 MW; 6869BA673438B628 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 146 VAEAEKAM 154  
 DB 137 VAEAEKAM 145  
 CC  
 CC RESULT 37  
 CC FABZ YERPS  
 CC ID FABZ YERPS STANDARD; PRT; 176 AA.  
 CC AC Q667K0;  
 CC DT 10-MAY-2005 (Rel. 47, Created)  
 CC DT 10-MAY-2005 (Rel. 47, Last sequence update)  
 CC DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 CC DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)  
 CC DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 CC GN Name=fabZ; OrderedLocusNames=YPTB2992;  
 CC OS Yersinia pseudotuberculosis.  
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC OC Enterobacteriaceae; Yersinia.  
 CC OX NCBI\_TaxID=633;  
 CC RN [1]  
 CC RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 CC RC STRAIN=IP32953 / Serotype I;  
 CC RX PubMed=15358858; DOI=10.1073/pnas.0404012101;  
 CC RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,  
 CC RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,  
 CC RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,  
 CC RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,  
 CC RA Derbise A., Hauser L.J., Garcia E.;  
 CC RT "Insights into the evolution of Yersinia pestis through whole-genome  
 CC comparison with Yersinia pseudotuberculosis";  
 CC Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).  
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL; BX936398; CAH22230.1; -; Genomic\_DNA.  
 CC HAMAP; MF 00406; -; 1.  
 CC InterPro; IPR010084; FabZ.  
 CC TIGRFAMs; TIGR01750; fabz; 1.  
 CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
 KW ACT SITE 54 By similarity.  
 FT ACT SITE 54  
 SQ SEQUENCE 176 AA; 19631 MW; 3A1602C0A93DD2E2 CRC64;  
 CC  
 CC Query Match 5.7%; Score 9; DB 1; Length 176;  
 CC Best Local Similarity 100.0%; Pred. No. 2;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 62 KPFPFGVLI 70  
 CC DB 58 KPFPFGVLI 66  
 CC  
 CC RESULT 38  
 CC FABZ YEREN  
 CC ID FABZ YEREN STANDARD; PRT; 178 AA.  
 CC AC P32205;  
 CC DT 01-OCT-1993 (Rel. 27, Created)  
 CC DT 15-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 CC DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)  
 CC DE ((3R)-hydroxymyristoyl ACP dehydratase).  
 CC GN Name=fabZ;





RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
 RL Nature 415:497-502(2002).  
 DR ENBL; AF464661; CAD14608.1; -; Genomic\_DNA.  
 DR GO; GO:0047794; F:cyclohexadienyl dehydrogenase activity; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0004665; F:prephenate dehydrogenase (NADP+) activity; IEA.  
 DR GO; GO:0008977; F:prephenate dehydrogenase activity; IEA.  
 DR GO; GO:0006571; P:tyrosine biosynthesis; IEA.  
 DR InterPro; IPR008235; PDH\_fb\_k\_insens.  
 DR InterPro; IPR003099; Prephen\_dehydrog.  
 DR Pfam; PF02153; PDH\_1.  
 DR PIRSF; PIRSF06786; PDH\_fb\_k\_insens; 1.  
 KW Complete proteome; Oxidoreductase.  
 SQ SEQUENCE 298 AA; 31444 MW; 8BD3F7C6B00683BD CRC64;

Query Match 5.7%; Score 9; DB 2; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 LKAMIAERD 159  
 |||||  
 Db 261 LKAMIAERD 269

RESULT 41  
 Q4YGS7 PLABE PRELIMINARY; PRT; 51 AA.  
 AC Q4YGS7;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN ORFNames=PB403260.00.0;  
 OS Plasmodium berghei.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5821;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Hall N., Karas M., Raine J.D., Carlton J.M., Kooij T.W.A.,  
 RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,  
 RA James K., Rutherford L., Harris B., Harris D., Churcher C.,  
 RA Quall M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,  
 RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,  
 RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;  
 RT "A comprehensive survey of the Plasmodium life cycle by genomic,  
 RT transcriptomic, and proteomic analyses.";  
 RL Science 307:82-86(2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAI01005230; CAI02790.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 51 AA; 6090 MW; A1187FBD3B43D4A CRC64;

Query Match 5.0%; Score 8; DB 2; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 7.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 YHLEVLKH 128  
 |||||  
 Db 11 YHLEVLKH 18

RESULT 42  
 Q90X75 GILMI PRELIMINARY; PRT; 94 AA.  
 ID Q90X75;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ABC A transporter (Fragment).  
 OS Gillichthys mirabilis (Long-jawed mudsucker).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Gobioidae;  
 OC Gobiidae; Gillichthys.  
 OX NCBI\_TaxID=8222;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Kultz D., Avila K.;  
 RL Submitted (SBP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF419321; AAL14706.1; -; mRNA.  
 FT NON\_TER 1  
 FT NON\_TER 94  
 SQ SEQUENCE 94 AA; 11206 MW; A52A5EFA04E984AE CRC64;

Query Match 5.0%; Score 8; DB 2; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 IIELOANK 40  
 |||||  
 Db 64 IIELOANK 71

RESULT 43  
 Q91404 XENLA PRELIMINARY; PRT; 105 AA.  
 AC Q91404;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE XK endo B homolog protein (Fragment).  
 GN Name=XK endo B homolog;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA MEDLINE=95290493; PubMed=7772598; DOI=10.1016/0167-4781(95)00049-M;  
 RA Adati N., Ito T., Koga C., Kito K., Sakaki Y., Shikawa K.;  
 RT "Differential display analysis of gene expression in developing  
 RT embryos of Xenopus laevis.";  
 RL Biochim. Biophys. Acta 1262:43-51(1995).  
 CC -!- SIMILARITY: Belongs to the intermediate filament family.  
 DR EMBL; S78089; AAB34522.1; -; mRNA.  
 DR PIR; S55691; S55691.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR001664; IF.  
 DR Pfam; PF00038; Filament; 1.  
 DR PROSITE; PS00226; IF; UNKNOWN\_1.  
 KW Intermediate filament.  
 FT NON\_TER 105  
 FT NON\_TER 105  
 SQ SEQUENCE 105 AA; 11896 MW; CDD924315223514A CRC64;

Query Match 5.0%; Score 8; DB 2; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 VDCKVVAE 148  
 |||||  
 Db 89 VDCKVVAE 96

RESULT 44  
 FABZ THEME STANDARD; PRT; 137 AA.  
 ID FABZ THEME  
 AC Q9WZQ8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

```

DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=TW0801;
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L.A., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,
RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE001747; AAD35883.1; -; Genomic_DNA.
DR FIC; B72335; B72335.
DR TIGR; TW0801; -.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 46 46 By similarity.
SQ SEQUENCE 137 AA; 15300 MW; ED2FE248F02299D0 CRC64;
-----
Query Match 5.0%; Score 8; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 66 PGVLIVEG 73
DB 54 PGVLIVEG 61
-----
RESULT 45
FABZ SYMTH STANDARD; PRT; 142 AA.
AC O67T91;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=STH117;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=T / IAM 14863;
RX PubMed=15383646; DOI=10.1093/nar/gkh830;
RA Ueda K., Yamashita A., Ighikawa J., Shimada M., Watsui T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Genome sequence of Symbiobacterium thermophilum, an uncultivable
RT bacterium that depends on microbial commensalism.";
RL Nucleic Acids Res. 32:4937-4944(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----

```

```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AP006840; BAD39102.1; -; Genomic_DNA.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 41 41 By similarity.
SQ SEQUENCE 142 AA; 15449 MW; 06C314FD13289981 CRC64;
-----
Query Match 5.0%; Score 8; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 63 PIFPGVLI 70
DB 46 PIFPGVLI 53
-----
RESULT 46
QSE3F0_VIBF1 PRELIMINARY; PRT; 150 AA.
AC QSE3F0;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
DE (EC 4.2.1.-)).
GN OrderedLocusNames=VF1951;
OS Vibrio fischeri (strain ATCC 700601 / BS114).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=312309;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15703294; DOI=10.1073/pnas.0409900102;
RA Ruby E.G., Urbanowski M., Campbell J., Dunn A., Faini M., Gunsalus R.,
RA Lostro P., Lupp C., McCann J., Millikan D., Schaefer A., Stabb E.,
RA Stevens A., Visick K., Whistler C., Greenberg E.P.;
RT "Complete genome sequence of Vibrio fischeri: a symbiotic bacterium
RT with pathogenic congeners.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:3004-3009(2005).
DR EMBL; CP000020; AAW86446.1; -; Genomic_DNA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 150 AA; 17033 MW; 9C75F3B016854B1D CRC64;
-----
Query Match 5.0%; Score 8; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 63 PIFPGVLI 70
DB 59 PIFPGVLI 66
-----
RESULT 47
FABZ_IDILO STANDARD; PRT; 151 AA.
AC QSR022;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE Name=fabZ; OrderedLocusNames=IL0834;
GN Idiomarina loihiensis.
OS

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
OC Idionarinaceae; Idiomarina.  
OX NCBI\_TaxID=135577;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=L2-TR / DSM 15497 / ATCC BAA-735;  
RX PubMed=15596722; DOI=10.1073/pnas.0407638102;  
RA Hou S., Saw J.H., Lee K.S., Freitas T.A., Belisle C., Kawarabayasi Y.,  
RA Donalchenko S.P., Pikina A., Galperin M.Y., Koonin E.V., Makarova K.S.,  
RA Omelchenko M.V., Sorokin A., Wolf Y.I., Li Q.X., Keum Y.S.,  
RA Campbell S., Denery J., Aizawa S.-I., Shibata S., Malahoff A.,  
RA Alam M.;  
RT "Genome sequence of the deep-sea gamma-proteobacterium Idiomarina  
RT loihensis reveals amino acid fermentation as a source of carbon and  
RT energy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:18036-18041(2004).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC EMBL; AE017340; AA081674.1; -; Genomic\_DNA.  
DR HAMAP; MF 00406; -; 1.  
DR InterPro; IPR010084; FabZ.  
DR TIGRFAMs; TIGR01750; fabz; 1.  
DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
KW ACT SITE 54 54 By similarity.  
FT ACT\_SITE 54 54 By similarity.  
SQ SEQUENCE 151 AA; 16991 MW; 784644D72C98C48A CRC64;  
  
Query Match 5.0%; Score 8; DB 1; Length 151;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 62 KPFFPGVL 69  
Db 58 KPFFPGVL 65  
|||||  
|||||  
  
RESULT 48  
FABZ PASMU STANDARD; PRT; 152 AA.  
AC Q9CJ99;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))  
DE ((3R)-hydroxymyristoyl ACP dehydratase).  
GN Names=fabZ; OrderedLocusNames=PM1995;  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Pm70;  
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC EMBL; AB006236; AA04079.1; -; Genomic\_DNA.  
DR HAMAP; MF 00406; -; 1.  
DR InterPro; IPR010084; FabZ.  
DR TIGRFAMs; TIGR01750; fabz; 1.  
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT ACT\_SITE 57 57 By similarity.  
SQ SEQUENCE 152 AA; 17291 MW; 7C38DDBSAFCB927 CRC64;  
  
Query Match 5.0%; Score 8; DB 1; Length 152;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 63 PIFPGVLI 70  
Db 62 PIFPGVLI 69  
|||||  
|||||  
  
RESULT 49  
FABZ RHIME STANDARD; PRT; 154 AA.  
AC Q92Q46;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))  
DE ((3R)-hydroxymyristoyl ACP dehydratase).  
GN Name=fabZ; OrderedLocusNames=R01504; ORFNames=SMC02092;  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=1021;  
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;  
RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
RT "Analysis of the chromosome sequence of the legume symbiont  
RT Sinorhizobium meliloti strain 1021.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC EMBL; AL591787; CAC46083.1; -; Genomic\_DNA.  
DR HAMAP; MF 00406; -; 1.  
DR InterPro; IPR010084; fabZ.  
DR TIGRFAMs; TIGR01750; fabz; 1.  
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.  
FT ACT\_SITE 57 57 By similarity.  
SQ SEQUENCE 154 AA; 16994 MW; 6E1BB0EA6F8CDF40 CRC64;  
  
Query Match 5.0%; Score 8; DB 1; Length 154;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 28 LLVDRIIE 35  
Db 27 LLVDRIIE 34  
|||||  
|||||  
  
RESULT 50  
Q9TIF3\_9CAUD



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:38:06 ; Search time 23 Seconds  
(without alignments)  
571.541 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHNLSQFFIEHILQI.....QVDGKVAEAEUKAMIAERD 159

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size : 8

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/prodata/1/iaa/5\_COMB.pep.\*
- 2: /cgn2\_6/prodata/1/iaa/6\_COMB.pep.\*
- 3: /cgn2\_6/prodata/1/iaa/H\_COMB.pep.\*
- 4: /cgn2\_6/prodata/1/iaa/PCITUS\_COMB.pep.\*
- 5: /cgn2\_6/prodata/1/iaa/RE\_COMB.pep.\*
- 6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	7.5	140	US-09-196-388-2	Sequence 2, Appli
2	12	7.5	140	US-09-583-110-3906	Sequence 3906, Ap
3	12	7.5	146	US-09-107-433-2892	Sequence 2892, Ap
4	12	7.5	160	US-10-089-019-18	Sequence 18, Appl
5	9	5.7	162	US-09-543-681A-6627	Sequence 6627, Ap

ALIGNMENTS

RESULT 1  
US-09-196-388-2

; Sequence 2, Application US/09196388

; Patent No. 6277595

; GENERAL INFORMATION:

; APPLICANT: Warren, Patrick

; APPLICANT: Konstantinidis, Alex

; APPLICANT: Russell, Robert

; TITLE OF INVENTION: fabz

; FILE REFERENCE: GM10182

; CURRENT APPLICATION NUMBER: US/09/196,388

; CURRENT FILING DATE: 1998-11-19

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 140

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae  
US-09-196-388-2

Query Match 7.5%; Score 12; DB 2; Length 140;

Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LPHRYPMLLVDR 32

Db 11 LPHRYPMLLVDR 22

RESULT 2

US-09-583-110-3906

; Sequence 3906, Application US/09583110

; Patent No. 669703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; FILE REFERENCE: PATH00-07A

; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 3906

; LENGTH: 140

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-583-110-3906

Query Match

Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LPHRYPMLLVDR 32

Db 11 LPHRYPMLLVDR 22

RESULT 3

US-09-107-433-2892

; Sequence 2892, Application US/09107433

; Patent No. 6800744

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: <Unknown>

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: <Unknown>

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,433

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/ 085131

; FILING DATE: May 12, 1998

; APPLICATION NUMBER: 60/051553  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 2892:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 146 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...146  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2892:  
US-09-107-433-2892

Query Match 7.5%; Score 12; DB 2; Length 146;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
||| |||||  
Db 17 LPHRYPMLLVDR 28

## RESULT 4

US-10-089-019-18  
; Sequence 18, Application US/10089019  
; Patent No. 6951729  
; GENERAL INFORMATION:  
; APPLICANT: DEWOLF, WALTER E. JR  
; APPLICANT: KALLENDER, HOWARD  
; APPLICANT: LONSDALE, JOHN T.  
; TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID  
; TITLE OF INVENTION: SYNTHESIS PATHWAY REAGENTS  
; FILE REFERENCE: GM50068  
; CURRENT APPLICATION NUMBER: US/10/089,019  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER: PCT/US00/29451  
; PRIOR FILING DATE: 2000-10-26  
; PRIOR APPLICATION NUMBER: 60/161,775  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-089-019-18

Query Match 7.5%; Score 12; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
||| |||||  
Db 31 LPHRYPMLLVDR 42

## RESULT 5

US-09-543-681A-6627  
; Sequence 6627, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6627  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6627

Query Match 5.7%; Score 9; DB 2; Length 162;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPGVLI 70  
||| |||||  
Db 69 KPIFPGVLI 77

Search completed: January 10, 2006, 13:42:09  
Job time : 23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:38:22 ; Search time 61 Seconds

(without alignments)  
1089,096 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHNLSQPFIEHILQI.....QVDGKVVAELKAMIAERD 159

Scoring table: OLIGO /

Gapop 60.0 , Gapext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size : 6

Total number of hits satisfying chosen parameters: 43

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 500 summaries

Database : Published Applications\_AA\_Main:\*  
1: /cgn2\_6/prodata1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata1/pubpaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/prodata1/pubpaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/prodata1/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata1/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/prodata1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	100.0	159	4	US-10-282-122A-58982
2	159	100.0	159	5	US-10-662-126-37
3	159	100.0	159	5	US-10-965-006-11
4	118	74.2	159	3	US-09-815-242-11601
5	118	74.2	159	4	US-10-335-977-6975
6	38	23.9	60	4	US-10-335-977-6974
7	12	7.5	139	4	US-10-282-122A-74705
8	12	7.5	140	3	US-09-815-242-13616
9	12	7.5	140	3	US-09-932-702-2
10	12	7.5	140	4	US-10-282-122A-73809
11	12	7.5	140	5	US-10-472-928-692
12	12	7.5	140	5	US-10-965-006-13
13	12	7.5	146	5	US-10-617-320-2892
14	11	6.9	141	4	US-10-282-122A-51842
15	10	6.3	146	4	US-10-282-122A-54269
16	9	5.7	102	4	US-10-282-122A-55351
17	9	5.7	102	4	US-10-282-122A-55353
18	9	5.7	146	5	US-10-650-274-133
19	9	5.7	151	3	US-09-815-242-10045
20	9	5.7	151	4	US-10-369-493-707
21	9	5.7	151	4	US-10-282-122A-42949
22	9	5.7	151	4	US-10-282-122A-51041
23	9	5.7	151	5	US-10-965-006-10
24	9	5.7	157	4	US-10-369-493-207
25	9	5.7	172	4	US-10-369-493-21205
26	9	5.7	172	4	US-10-282-122A-68573
27	9	5.7	181	4	US-10-282-122A-77967

28	8	5.0	133	3	US-09-738-769A-4	Sequence 4, Appli
29	8	5.0	137	4	US-10-369-493-2971	Sequence 2971, Ap
30	8	5.0	139	4	US-10-282-122A-57925	Sequence 57925, A
31	8	5.0	152	4	US-10-282-122A-67466	Sequence 67466, A
32	8	5.0	154	4	US-10-369-493-19450	Sequence 19450, A
33	8	5.0	222	4	US-10-062-254-88	Sequence 88, Appl
34	8	5.0	222	4	US-10-767-701-35176	Sequence 35176, A
35	8	5.0	437	4	US-10-425-114-46464	Sequence 46464, A
36	8	5.0	478	4	US-10-369-493-2615	Sequence 2615, Ap
37	8	5.0	561	4	US-10-437-963-203850	Sequence 203850,
38	8	5.0	562	4	US-10-425-115-231422	Sequence 231422,
39	8	5.0	565	4	US-10-062-254-92	Sequence 92, Appl
40	8	5.0	565	4	US-10-425-115-300838	Sequence 300838,
41	8	5.0	590	4	US-10-437-963-198210	Sequence 198210,
42	8	5.0	684	4	US-10-156-761-10463	Sequence 10463, A
43	8	5.0	910	4	US-10-369-493-21509	Sequence 21509, A

## ALIGNMENTS

### RESULT 1

US-10-282-122A-58982

; Sequence 58982, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 58982

; LENGTH: 159

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-10-282-122A-58982

Query Match 100.0%; Score 159; DB 4; Length 159;  
Best Local Similarity 100.0%; Pred. No. 9e-155;

Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDVFNHGHP 60  
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDVFNHGHP 60  
QY 61 NKPIPPGVLIVVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIPPGVLIVVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
QY 121 YHLEVLKHKGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159  
DB 121 YHLEVLKHKGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159

RESULT 2  
US-10-662-126-37  
; Sequence 37, Application US/10662126  
; Publication No. US20050063987A1  
; GENERAL INFORMATION:  
; APPLICANT: Knapp, Bernhard  
; APPLICANT: Hundt, Erika  
; APPLICANT: Schmidt, Karl-Heinz  
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of  
; Helicobacter Pylori, Their Preparation and Use  
; FILE REFERENCE: CHIR-0340  
; CURRENT APPLICATION NUMBER: US/10/662,126  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: 09/230,158  
; PRIOR FILING DATE: 1999-01-19  
; PRIOR APPLICATION NUMBER: PCT/IB97/00981  
; PRIOR FILING DATE: 1997-07-25  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 37  
; TYPE: PRT  
; LENGTH: 159  
; ORGANISM: Helicobacter pylori  
US-10-662-126-37

Query Match 100.0%; Score 159; DB 5; Length 159;  
Best Local Similarity 100.0%; Pred. No. 9e-155;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDVFNHGHP 60  
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDVFNHGHP 60  
QY 61 NKPIPPGVLIVVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIPPGVLIVVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
QY 121 YHLEVLKHKGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159  
DB 121 YHLEVLKHKGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159

RESULT 3  
US-10-965-006-11  
; Sequence 11, Application US/10965006  
; Publication No. US20050221462A1  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, ALED  
; APPLICANT: DHARAMSI, AKIL  
; APPLICANT: VEDADI, MASOUD  
; APPLICANT: DOMAGALA, MEGAN  
; APPLICANT: KIMBER, MATTHEW  
; APPLICANT: VALLEE, FRANCOIS  
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA  
; FILE REFERENCE: IPT-257.01  
; CURRENT APPLICATION NUMBER: US/10/965,006  
; CURRENT FILING DATE: 2004-10-14  
; PRIOR APPLICATION NUMBER: PCT/CA03/00560

; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: 60/373,321  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 3.3  
; SEQ ID NO 11  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-10-965-006-11

Query Match 100.0%; Score 159; DB 5; Length 159;  
Best Local Similarity 100.0%; Pred. No. 9e-155;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDVFNHGHP 60  
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDVFNHGHP 60  
QY 61 NKPIPPGVLIVVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
DB 61 NKPIPPGVLIVVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120  
QY 121 YHLEVLKHKGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159  
DB 121 YHLEVLKHKGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159

RESULT 4  
US-09-815-242-11601  
; Sequence 11601, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11601  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-815-242-11601

Query Match 74.2%; Score 118; DB 3; Length 159;  
Best Local Similarity 100.0%; Pred. No. 1.1e-112;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 KIVAYKNITFNEDVFNHGHPNKPFPFPGVLIVVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVY 100

Db 41 KIVAYKNITFNEGVNFKKIPFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTVY 100  
QY 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMWQVGGTAQVDGKVVAAELKAMIAER 158  
Db 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMWQVGGTAQVDGKVVAAELKAMIAER 158  
RESULT 5  
US-10-335-977-6975  
; Sequence 6975, Application US/10335977  
; Publication No. US20040052799A1  
; GENERAL INFORMATION:  
; APPLICANT: DOUGLAS SMITH et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; RELATING TO HELICOBACTER PYLORI FOR  
; DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 10031  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows NT 4.0  
; SOFTWARE: UNIX  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/335,977  
; FILING DATE: 30-Dec-2002  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/993,002  
; FILING DATE: 17-DEC-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: GTN-018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 6975:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 159 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...159  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6975:  
US-10-335-977-6975  
Query Match 74.2%; Score 118; DB 4; Length 159;  
Best Local Similarity 100.0%; Pred. No. 1.1e-112;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 41 KIVAYKNITFNEGVNFKKIPFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTVY 100  
Db 41 KIVAYKNITFNEGVNFKKIPFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTVY 100  
QY 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMWQVGGTAQVDGKVVAAELKAMIAER 158  
Db 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMWQVGGTAQVDGKVVAAELKAMIAER 158  
RESULT 6  
US-10-335-977-6974

; Sequence 6974, Application US/10335977  
; Publication No. US20040052799A1  
; GENERAL INFORMATION:  
; APPLICANT: DOUGLAS SMITH et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; RELATING TO HELICOBACTER PYLORI FOR  
; DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 10031  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows NT 4.0  
; SOFTWARE: UNIX  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/335,977  
; FILING DATE: 30-Dec-2002  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/993,002  
; FILING DATE: 17-DEC-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: GTN-018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 6974:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 60 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...60  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6974:  
US-10-335-977-6974  
Query Match 23.9%; Score 38; DB 4; Length 60;  
Best Local Similarity 100.0%; Pred. No. 6.3e-31;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 58 HFPNKPFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAK 95  
Db 5 HFPNKPFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAK 42  
RESULT 7  
US-10-282-122A-74705  
; Sequence 74705, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.

```
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 74705
/ LENGTH: 139
/ TYPE: PRT
/ ORGANISM: Streptococcus pyogenes
US-10-282-122A-74705

Query Match          7.5%; Score 12; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
   |||||
DB 10 LPHRYPMLLVDR 21

RESULT 8
US-09-815-242-13616
/ Sequence 13616, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
```

```
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 13616
/ LENGTH: 140
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-09-815-242-13616

Query Match          7.5%; Score 12; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
   |||||
DB 11 LPHRYPMLLVDR 22

RESULT 9
US-09-932-702-2
/ Sequence 2, Application US/09932702
/ Patent No. US20020119520A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Patrick
/ APPLICANT: Konstantinidis, Alex
/ APPLICANT: Russell, Robert
/ TITLE OF INVENTION: fabZ
/ FILE REFERENCE: GM10182
/ CURRENT APPLICATION NUMBER: US/09/932,702
/ CURRENT FILING DATE: 2001-08-17
/ PRIOR APPLICATION NUMBER: 09/196,388
/ PRIOR FILING DATE: 1998-11-19
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 140
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-09-932-702-2

Query Match          7.5%; Score 12; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
   |||||
DB 11 LPHRYPMLLVDR 22

RESULT 10
US-10-282-122A-73809
/ Sequence 73809, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
```

; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 73809  
; TYPE: PRT  
; LENGTH: 140  
; ORGANISM: Streptococcus pneumoniae  
US-10-282-122A-73809

Query Match 7.5%; Score 12; DB 4; Length 140;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
Db 11 LPHRYPMLLVDR 22

RESULT 11  
US-10-472-928-692  
; Sequence 692, Application US/10472928  
; Publication No. US20050020813A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH  
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026926W0  
; CURRENT APPLICATION NUMBER: US/10/472,928  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: GB-0107658.7  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 4979  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 692  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; OTHER INFORMATION: (3R)-hydroxymyristoyl-(acyl-carrier-protein) dehydratase (fabZ)  
; OTHER INFORMATION: Cellular location: cytoplasm  
; OTHER INFORMATION: Similar to strain R6 sequence 15902428 (7.E-75)  
US-10-472-928-692

Query Match 7.5%; Score 12; DB 5; Length 140;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
Db 11 LPHRYPMLLVDR 22

RESULT 12  
US-10-965-006-13  
; Sequence 13, Application US/10965006

; Publication No. US20050221462A1  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, ALED  
; APPLICANT: DHARAMSI, AKIL  
; APPLICANT: VEDADI, MASOUD  
; APPLICANT: DOMAGALA, MEGAN  
; APPLICANT: KIMBER, MATTHEW  
; APPLICANT: VALLEE, FRANCOIS  
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA  
; FILE REFERENCE: IPT-257.01  
; CURRENT APPLICATION NUMBER: US/10/965,006  
; CURRENT FILING DATE: 2004-10-14  
; PRIOR APPLICATION NUMBER: PCT/CA03/00560  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: 60/373,321  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 3.3  
; SEQ ID NO 13  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-965-006-13

Query Match 7.5%; Score 12; DB 5; Length 140;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32  
Db 11 LPHRYPMLLVDR 22

RESULT 13  
US-10-617-320-2892  
; Sequence 2892, Application US/10617320  
; Publication No. US20050136404A1  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID  
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE  
; THERAPEUTICS  
; NUMBER OF SEQUENCES: 5206  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: <Unknown>  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: <Unknown>  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/617,320  
; FILING DATE: 10-Jul-2003  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,433  
; FILING DATE: 30-Jun-1998  
; APPLICATION NUMBER: 60/ 085131  
; FILING DATE: May 12, 1998  
; APPLICATION NUMBER: 60/051553  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 2892:

```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 146 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: YES
;   ORIGINAL SOURCE:
;     ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;     NAME/KEY: misc feature
;     LOCATION: (8) LOCATION 1...146
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2892:
US-10-617-320-2892

```

```

Query Match          7.5%; Score 12; DB 5; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 21 LPHRYPMLLVDR 32
    |||||
Db 17 LPHRYPMLLVDR 28

```

## RESULT 14

```

US-10-282-122A-51842
; Sequence 51842, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Liangsu
;   APPLICANT: Zamudio, Carlos
;   APPLICANT: Malone, Cheryl
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Karl
;   APPLICANT: Zyskind, Judith
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John
;   APPLICANT: Carr, Grant
;   APPLICANT: Yamamoto, Robert
;   APPLICANT: Forsyth, R.
;   APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51842
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51842

```

```

Query Match          6.9%; Score 11; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0069;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 22 PHRYPMLLVDR 32
    |||||
Db 13 PHRYPMLLVDR 23

```

## RESULT 15

```

US-10-282-122A-54269
; Sequence 54269, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Liangsu
;   APPLICANT: Zamudio, Carlos
;   APPLICANT: Malone, Cheryl
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Karl
;   APPLICANT: Zyskind, Judith
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John
;   APPLICANT: Carr, Grant
;   APPLICANT: Yamamoto, Robert
;   APPLICANT: Forsyth, R.
;   APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54269
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-10-282-122A-54269

```

```

Query Match          6.3%; Score 10; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 146 VAEAEKXAMI 155
    |||||
Db 134 VAEAEKXAMI 143

```

## RESULT 16

```

US-10-282-122A-55351
; Sequence 55351, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

```

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 55351  
LENGTH: 102  
TYPE: PRT  
ORGANISM: Enterobacter cloacae  
US-10-282-122A-55351

Query Match 5.7%; Score 9; DB 4; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.59;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPFPFGVLI 70  
Db 74 KPFPFGVLI 82

RESULT 17  
US-10-282-122A-55353  
Sequence 55353, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 55353  
LENGTH: 102  
TYPE: PRT  
ORGANISM: Enterobacter cloacae  
US-10-282-122A-55353

Query Match 5.7%; Score 9; DB 4; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.59;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPFPFGVLI 70  
Db 74 KPFPFGVLI 82

RESULT 18  
US-10-650-274-153  
Sequence 153, Application US/10650274  
Publication No. US20050202437A1  
GENERAL INFORMATION:  
APPLICANT: GLENN, MATTHEW  
APPLICANT: HAVUKKALA, ILKKA J  
APPLICANT: LOBBERS, MARK WILLIAM  
APPLICANT: DEKKER, JAMES  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES,  
TITLE OF INVENTION: MATERIALS INCORPORATING THEM, AND METHODS FOR USING  
TITLE OF INVENTION: THEM.  
FILE REFERENCE: 11000.1073  
CURRENT APPLICATION NUMBER: US/10/650,274  
CURRENT FILING DATE: 2003-08-28  
NUMBER OF SEQ ID NOS: 187  
SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0  
SEQ ID NO 153  
LENGTH: 146  
TYPE: PRT  
ORGANISM: LACTOBACILLUS RHAMNOSUS  
US-10-650-274-153

Query Match 5.7%; Score 9; DB 5; Length 146;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ILPHRYPML 28  
Db 16 ILPHRYPML 24

RESULT 19



```
US-09-815-242-10045
; Sequence 10045, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10045
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10045

Query Match          5.7%; Score 9; DB 3; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      62 KPIFPGLI 70
Db      58 KPIFPGLI 66

RESULT 20
US-10-369-493-707
; Sequence 707, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 707
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-707

Query Match          5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      62 KPIFPGLI 70
Db      58 KPIFPGLI 66

US-09-815-242-10045
; Sequence 10045, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10045
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10045

Query Match          5.7%; Score 9; DB 3; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      62 KPIFPGLI 70
Db      58 KPIFPGLI 66

RESULT 20
US-10-369-493-707
; Sequence 707, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 707
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-707

Query Match          5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      62 KPIFPGLI 70
Db      58 KPIFPGLI 66

US-10-282-122A-42949
; Sequence 42949, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42949
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-42949

Query Match          5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      62 KPIFPGLI 70
Db      58 KPIFPGLI 66

RESULT 22
US-10-282-122A-51041
; Sequence 51041, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 51041  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Bordetella pertussis  
US-10-282-122A-51041

Query Match 5.7%; Score 9; DB 4; Length 151;  
Best Local Similarity 100.0%; Pred. No. 0.83;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LPHRYPMLL 29  
Db 12 LPHRYPMLL 20

RESULT 23  
US-10-965-006-10  
Sequence 10, Application US/10965006  
Publication No. US20050221462A1  
GENERAL INFORMATION:  
APPLICANT: EDWARDS, ALED  
APPLICANT: DHARAMSI, AKIL  
APPLICANT: VEDADI, MASOUD  
APPLICANT: DOMAGALA, MEGAN  
APPLICANT: KIMBER, MATTHEW  
APPLICANT: VALLBER, FRANCOIS  
TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA  
FILE REFERENCE: IPT-257.01  
CURRENT APPLICATION NUMBER: US/10/965,006  
CURRENT FILING DATE: 2004-10-14  
PRIOR APPLICATION NUMBER: PCT/CA03/00560  
PRIOR FILING DATE: 2003-04-17  
PRIOR APPLICATION NUMBER: 60/373,321  
PRIOR FILING DATE: 2002-04-17  
NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 3.3  
SEQ ID NO 10  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-10-965-006-10

Query Match 5.7%; Score 9; DB 5; Length 151;  
Best Local Similarity 100.0%; Pred. No. 0.83;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70  
Db 58 KPIPPGVLI 66

RESULT 24  
US-10-369-493-207  
Sequence 207, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 207  
LENGTH: 157  
TYPE: PRT  
ORGANISM: Xenorhabdus nematophilus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(157)  
OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-207

Query Match 5.7%; Score 9; DB 4; Length 157;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70  
Db 64 KPIPPGVLI 72

RESULT 25  
US-10-369-493-21205  
Sequence 21205, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 21205  
LENGTH: 172

```

US-10-282-122A-68573
TYPE: PRT
ORGANISM: Xenorhabdus nematophilus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(172)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-21205
Query Match 5.7%; Score 9; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFFPGVLI 70
DB 79 KPIFFPGVLI 87

RESULT 26
US-10-282-122A-68573
; Sequence 68573, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68573
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X-any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (12)..(12)
; OTHER INFORMATION: X-any amino acid

US-10-282-122A-68573
TYPE: PRT
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(181)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-21205
Query Match 5.7%; Score 9; DB 4; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFFPGVLI 70
DB 58 KPIFFPGVLI 66

RESULT 28
US-10-282-122A-77967
; Sequence 77967, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77967
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Yersinia pestis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X-any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (12)..(12)
; OTHER INFORMATION: X-any amino acid

```

```
; Patent No. US20020055148A1
; GENERAL INFORMATION:
; APPLICANT: HONG, Kyongsu
; TITLE OF INVENTION: A CLONING METHOD FOR DNA FRAGMENTS USING ARBITRARILY PRIMED PCR
; FILE REFERENCE: P20153
; CURRENT APPLICATION NUMBER: US/09/738,769A
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 064432/99
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 09/519,581
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Streptococcus zooepidemicus
US-09-738-769A-4

Query Match          5.0%; Score 8; DB 3; Length 133;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      26 PMLLVDR1 33
      |||||
Db      9 PMLLVDR1 16

RESULT 29
US-10-369-493-2971
; Sequence 2971, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2971
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2971

Query Match          5.0%; Score 8; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      66 PGVLIVEG 73
      |||||
Db      54 PGVLIVEG 61

RESULT 30
US-10-282-122A-57925
; Sequence 57925, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

```
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57925
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57925

Query Match          5.0%; Score 8; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      22 PHRYPMLL 29
      |||||
Db      11 PHRYPMLL 18

RESULT 31
US-10-282-122A-67466
; Sequence 67466, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining prior application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 67466  
 ; LENGTH: 152  
 ; TYPE: PRT  
 ; ORGANISM: Pasteurella multocida  
 US-10-282-122A-67466

Query Match 5.0%; Score 8; DB 4; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 PIPPGVLI 70  
 |||||  
 DB 62 PIPPGVLI 69

RESULT 32  
 US-10-369-493-19450  
 ; Sequence 19450, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(S2052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; PRIOR FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 19450  
 ; LENGTH: 154  
 ; TYPE: PRT  
 ; ORGANISM: Myxococcus xanthus  
 US-10-369-493-19450

Query Match 5.0%; Score 8; DB 4; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPML 28  
 |||||  
 DB 17 LPHRYPML 24

RESULT 33  
 US-10-062-254-88  
 ; Sequence 88, Application US/10062254  
 ; Publication No. US2002013882A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Caboon, Edgar B  
 ; APPLICANT: Cahoon, Rebecca E

; APPLICANT: Falco, Saverio Carl  
 ; APPLICANT: Fang, Yiwen  
 ; APPLICANT: Hantke, Sabine S.  
 ; APPLICANT: Lee, Jian-Ming  
 ; APPLICANT: Li, Zhongsen  
 ; APPLICANT: Miao, Guo-Hua  
 ; APPLICANT: Morgante, Michele  
 ; APPLICANT: Niu, Xiping  
 ; APPLICANT: Odell, Joan  
 ; APPLICANT: Rafalski, Antoni  
 ; APPLICANT: Sakai, Hajime  
 ; APPLICANT: Zheng, Peizhong  
 ; APPLICANT: Zhu, Qun

; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism  
 ; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/062,254  
 ; CURRENT FILING DATE: 2002-02-01  
 ; PRIOR APPLICATION NUMBER: 09/630,346  
 ; PRIOR FILING DATE: 2000-07-28  
 ; PRIOR APPLICATION NUMBER: 60/146511  
 ; PRIOR FILING DATE: 1999-07-30  
 ; PRIOR APPLICATION NUMBER: 60/156006  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: 60/156899  
 ; PRIOR FILING DATE: 1999-09-30  
 ; PRIOR APPLICATION NUMBER: 60/157287  
 ; PRIOR FILING DATE: 1999-10-01  
 ; PRIOR APPLICATION NUMBER: 60/169767  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 60/171054  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: 60/172958  
 ; PRIOR FILING DATE: 1999-12-21  
 ; PRIOR APPLICATION NUMBER: 60/171515  
 ; PRIOR FILING DATE: 1999-12-22  
 ; PRIOR APPLICATION NUMBER: 60/173535  
 ; PRIOR FILING DATE: 1999-12-29  
 ; NUMBER OF SEQ ID NOS: 375  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 88  
 ; LENGTH: 222  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (4)  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (9)  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (18)  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (26)  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (29)  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (52)  
 US-10-062-254-88

Query Match 5.0%; Score 8; DB 4; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95  
 |||||  
 DB 156 GFDPEIAK 163

RESULT 34  
 US-10-767-701-35176  
 ; Sequence 35176, Application US/10767701  
 ; Publication No. US20040172684A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 35176
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C60519_1.pep
US-10-767-701-35176

Query Match          5.0%; Score 8; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      88 GFDPEIAK 95
Db      95 GFDPEIAK 102

RESULT 35
US-10-425-114-46464
; Sequence 46464, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46464
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700209195_FLI.pep
US-10-425-114-46464

Query Match          5.0%; Score 8; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      88 GFDPEIAK 95
Db      160 GFDPEIAK 167

RESULT 36
US-10-369-493-2615
; Sequence 2615, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
```

```
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2615
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Synechocystis sp.
US-10-369-493-2615

Query Match          5.0%; Score 8; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      88 GFDPEIAK 95
Db      219 GFDPEIAK 236

RESULT 37
US-10-437-963-203850
; Sequence 203850, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203850
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98995C.1.pep
US-10-437-963-203850

Query Match          5.0%; Score 8; DB 4; Length 561;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      88 GFDPEIAK 95
Db      284 GFDPEIAK 291

RESULT 38
US-10-425-115-231422
; Sequence 231422, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231422
; LENGTH: 562
; TYPE: PRT
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_142653C.1.pep
US-10-425-115-231422

Query Match          5.0%; Score 8; DB 4; Length 562;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      285 GFDPEIAK 292

RESULT 39
US-10-062-254-92
; Sequence 92, Application US/10062254
; Publication No. US2002013882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 92
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Zea mays
US-10-062-254-92

Query Match          5.0%; Score 8; DB 4; Length 565;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      288 GFDPEIAK 295

RESULT 40
US-10-425-115-300838
; Sequence 300838, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 300838
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_37433C.1.pep
US-10-425-115-300838

Query Match          5.0%; Score 8; DB 4; Length 565;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      288 GFDPEIAK 295

RESULT 41
US-10-437-963-198210
; Sequence 198210, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198210
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93893C.1.pep
US-10-437-963-198210

Query Match          5.0%; Score 8; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      290 GFDPEIAK 297

RESULT 42
US-10-156-761-10463
; Sequence 10463, Application US/10156761
```

```
RESULT 40
US-10-425-115-300838
; Sequence 300838, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 300838
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_37433C.1.pep
US-10-425-115-300838

Query Match          5.0%; Score 8; DB 4; Length 565;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      288 GFDPEIAK 295

RESULT 41
US-10-437-963-198210
; Sequence 198210, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198210
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93893C.1.pep
US-10-437-963-198210

Query Match          5.0%; Score 8; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      290 GFDPEIAK 297

RESULT 42
US-10-156-761-10463
; Sequence 10463, Application US/10156761
```



; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 10463

; LENGTH: 684

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-10463

Query Match 5.0%; Score 8; DB 4; Length 684;

Best Local Similarity 100.0%; Pred. No. 34;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGFLAFTS 85

Db 293 GGFLAFTS 300

RESULT 43

US-10-369-493-21509

; Sequence 21509, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 21509

; LENGTH: 910

; TYPE: PRT

; ORGANISM: Methanobacterium thermoautotrophicum

US-10-369-493-21509

Query Match 5.0%; Score 8; DB 4; Length 910;

Best Local Similarity 100.0%; Pred. No. 43;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 AEAELKAM 154

Db 424 AEAELKAM 431

Search completed: January 10, 2006, 13:43:21

Job time : 61 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:39:52 ; Search time 9 Seconds  
(without alignments)  
150.054 Million cell updates/sec

Title: US-10-662-126-37  
Perfect score: 159  
Sequence: 1 MEQSHNLOSQFFIEHLQI.....QVDGKVVAEAEKAMIAERD 159

Scoring table: OLIGO  
Gapop: 60.0 , Gapext 60.0

Searched: 61141 seqs, 8493638 residues

Word size : 8

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

- Database : Published Applications\_AA\_New.\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----					

No matches found

Search completed: January 10, 2006, 13:43:40  
Job time : 9 secs